Individual Characteristics in Performance Appraisal

Fang Wang

A Thesis
in
the John Molson School of Business

Presented in Partial Fulfillment of the Requirements
for the Degree of Master of Science in Administration at
Concordia University
Montreal, Quebec, Canada
November 2010

© Fang Wang, 2010

CONCORDIA UNIVERSITY
School of Graduate Studies

This is to certify that the thesis prepared

By: Fang Wang

Entitled: Individual Characteristics in Performance Appraisal

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

Master of Science in 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:

Michele Paulin Chair

Linda Dyer Examiner

Marylène Gagné Examiner

Stéphane Brutus Supervisor

Approved by _____
Chair of Department or Graduate Program Director

_____
Dean of Faculty

Date November 22, 2010

ABSTRACT

Individual Characteristics in Performance Appraisal

Fang Wang

In the past two decades, studies about various formats used in performance appraisal have been focused on examining the format of graphic or numeric ratings, because of their advantages of being relatively objective and convenient to obtain quick and easy performance measurement. This paper emphasizes that solely quantitative data seems to be tenuous to satisfy a broad range of evaluation purposes and underscores the value of qualitative data, i.e. narrative comments in performance appraisal.

The purpose of this study was to examine the relationship between the personality traits of Conscientiousness, Extraversion and Agreeableness, and user's production of comments in evaluation. Through a study conducted among undergraduate students who used peer evaluation system, results revealed that the three characteristics had no effects on the amount, specificity or positivity of comments, and group size did not moderate the relationship between the personality traits and production of comments. Nevertheless, gender was found related to the specificity of comments. Findings and limitations, as well as practical implications, are discussed. More extensive research effort is called for about comments in performance appraisal.

Acknowledgements

First of all, I would like to express my sincere appreciation to my thesis supervisor, Dr. Brutus, for his continuous guidance and generous support throughout the process. Second, I would like to thank Dr. Dyer and Dr. Gagné for their helpful suggestions during my thesis writing, as well as the knowledge that I learned in their classes. Also, I would like to thank Amanda for her time working as the coder of my thesis.

Last, but not least, I would like to give my great appreciation to my husband and my parents, who have always loved and supported me.

TABLE OF CONTENTS

LIST OF FIGURES…….. vii

LIST OF TABLES…….. viii

Introduction……..... 1

Numeric Ratings in Performance Appraisals……. 1

Narrative Comments in Performance Appraisals……... 3

Theoretical Framework…….…….. 9

Big-Five Factors of Personality Traits……..... 9

Big-Five Factor Model and Performance Appraisal…….… 12

Purposes of Research……....… 14

Research Model and Hypotheses……. 14

Conscientiousness and Writing Comments…….. 15

Extraversion and Writing Comments……... 18

Agreeableness and Writing Comments……..... 21

Group Size and Writing Comments……..... 23

Research Methodology……..... 27

Sample……..... 27

Procedure…….. 28

Measure……..... 29

Independent Variables……..... 29

Dependent Variables……....… 30

Moderator…….. 34

Results……... 34

Discussion…… 37

Social Context of Peer Evaluation…… 39

Sample Limitations…….. 40
Quality of Peer Evaluation……...… 41
Future Research……..... 42
Theoretical Contribution and Practical Implications…… 46
References…… 50
Appendices…….….. 61
Appendix A — Tables……. 61
Appendix B — Participant Recruitment Letter…… 67
Appendix C — Consent Form……..... 68
Appendix D — Questionnaire……...….... 71
Appendix E — Coding Protocol……... 76

LIST OF FIGURES

Figure 1. Hypotheses 1, 2, 3, 4 and 5…… 25

Figure 2. Hypotheses 6a and 6b…… 26

Figure 3. Hypotheses 7a and 7b…… 26

LIST OF TABLES

Table 1. Intercorrelations Among Key Variables……. 61

Table 2. Hierarchical Regression Analyses with Amount of Comments as Dependent Variable, Conscientiousness as Predictor, and Group Size as Moderator……...……. 62

Table 3. Hierarchical Regression Analyses with Specificity of Comments as Dependent Variable, Conscientiousness as Predictor, and Group Size as Moderator…… 63

Table 4. Hierarchical Regression Analyses with Amount of Comments as Dependent Variable, Extraversion as Predictor, and Group Size as Moderator…… 64

Table 5. Hierarchical Regression Analyses with Specificity of Comments as Dependent Variable, Extraversion as Predictor, and Group Size as Moderator……. 65

Table 6. Hierarchical Regression Analyses with Positivity of Comments as Dependent Variable and Agreeableness as Predictor…….. 66

Introduction

Performance appraisal has been one of the most widely and frequently studied areas in industrial-organizational psychology and human resource management research, and one of the most frequently discussed dimensions of performance appraisal has been the formats that are employed. As "the vehicle" (Landy & Farr, 1980, p. 82) to deliver specific evaluation information, formats have essential influences on the value and accuracy of the information that will be provided to appraisees. In general, commonly used formats include graphic or numeric rating scales, open-ended, unstructured descriptive comments, ranking methods and etc.

This study will discuss the advantages and the disadvantages of numeric ratings and narrative comments and then examine the influences of certain personality traits on production of narrative comments.

Numeric Ratings in Performance Appraisals

During 1980s, while performance appraisal formats received some attention from practitioners, relevant scholarly research was limited (Bretz, Milkovich & Read, 1992). In the past two decades, studies about various formats used in performance appraisal have proliferated. However, most of them have been inclined to discussing the format of graphic or numeric ratings, by which individual performance is evaluated through assigning numbers that represent certain levels of criteria required for the job. In particular, Behavioral Observation Scales (BOS) and Behaviorally Anchored Rating Scales (BARS) are the two graphic rating scales that have been most popular in

organizations and most frequently compared by researchers, because of their advantages of being relatively objective and quantitative (Tziner & Kopelman, 2002). In a majority of human resource management literature, the term "performance evaluation" can almost be replaced by "performance rating", and it is common to find "evaluator" being called "rater" or "evaluatee" being called "ratee" in the literature.

It is understandable why quantitative data has been more desired by practitioners as well as researchers than qualitative data. First, quantitative data can avoid subjective coding possibility that is likely to threaten the reliability and validity of the results (Jackson & Trochim, 2002). Second, a quantitative survey requires much less time to complete and quantitative data demand much less effort to interpret, which implies much more time-saving and cost-saving evaluation and analysis processes. This appears to be especially important nowadays in the fast-changing and efficiency-oriented workplace. Last, but not least, social or political contexts within organizations, which can have significant impacts on the process and results of performance appraisal (Ferris, Munyon, Basik, & Buckley, 2008; Fried, Tiegs, & Bellamy, 1992; Johns, 2006; Judge & Ferris, 1993), also contribute to the popularity of quantitative format of performance appraisal. As evaluators, managers' political or social interests may overwhelm their concerns for an accurate performance evaluation (Fisher, 1989). To be more specific, in some cases, out of a purpose of maintaining good relationship or avoiding conflicts with their subordinates, supervisors may feel more comfortable conducting quantitative appraisal than writing formal evaluation information (Fried et al., 1992; Tziner & Kopelman, 2002). It should be a similar story when it comes to peer evaluation. Narrative comments usually contain the evaluator's personal opinions on the evaluatee's performance, personality or

quality, which may include negative opinions, and the evaluatee is likely to have a chance to read the comments afterwards. Therefore, out of a concern of potential damage to the interpersonal relationship in the company, employees may show avoidance of writing comments for their co-workers.

However, other than administrative needs such as promotions, pay adjustments or downsizing that ask for quick and easy performance measurement, which lead to dominance of quantitative data, performance appraisal also serves a number of other purposes in organizations, including employee feedback and development, training programs and personnel research (Sulsky & Keown, 1998). Quantitative data alone seems to be tenuous to satisfy such a broad range of evaluation purposes. Some researchers have brought up the necessity of an integrated approach of conducting performance appraisal in order to better serve the multiple needs of organizations (Bretz et al., 1992; Brutus, 2010; Ilgen, Barnes-Farrell, & Mckellin, 1993; Schweiger & Sumners, 1994; Sulsky & Keown, 1998).

Narrative Comments in Performance Appraisals

Qualitative data, usually in a form of narrative texts responding to open-ended questions in surveys, has its unique advantages. First, it provides a "rich description of respondent reality" (Jackson & Trochim, 2002, p. 307) without narrowing respondents' frames of reference during their information search in their mind and thus elicits diversity in responses. Second, it helps "explain or clarify quantitative findings" (Jackson & Trochim, 2002, p. 307). Numeric ratings in performance appraisal become more convincing and justifiable while underpinned by the evaluator's narrative comments

(Antonioni, 1996). This may as well help improve employees' fairness perceptions, which is identified by managers as the most important performance appraisal issue organizations face (Bretz et al., 1992). Third, quantitative data often appear to be tenuous to fully reflect diverse or changing organizational contexts, and as a complement, qualitative data can better establish contextual links between performance and appraisal by illuminating context effects (Johns, 2006). Fourth, especially in those performance appraisals conducted for purpose of employee feedback or development, qualitative data, i.e. written comments about specific work behaviors, can provide feedback information with more details for employees' reflection (Schweiger & Sumners, 1994). After all, the ultimate goals of performance appraisals for such purposes are not only to justify administrative decisions, but, more importantly, to provide an opportunity both for employees to reflect on their past performance and for managers to communicate their suggestions and future expectations regarding subordinates' performance improvement (Bretz et al., 1992).

Furthermore, in presence of numbers and texts in their performance appraisal, evaluatees tend to heed and respond more the latter, which suggests that quantitative and qualitative data do have different weights in recipients' mind (Antonioni, 1996; Delassio, 1998; Waldman, Atwater, & Antonioni, 1998). In their integration report of performance appraisal practices in US organizations, Bretz et al. (1992) found that it was common for "60% to 70% of an organization's workforce to be rated in the top two performance levels" (p. 333). In that case, when the average rated performance level in one organization is distributed at the top end, many employees may just ignore the performance level that he or she is assigned to, which is usually based on calculation of

rating numbers, but pay more attention to the specific feedback information relating to their certain job skills or performance. This could be one possible reason why appraisal recipients show more interest in written feedback comments for them than numeric ratings. In a theoretical approach, the speculation can be reconfirmed by Brutus' explanation (2010) by media richness theory (Daft & Lengel, 1984, 1986). Media richness can generally be defined as a medium's ability to communicate meanings. A rich medium can transmit messages more clearly and more effectively than a lean one. In organizational communications, capacity for personal focus is one of the indications of media richness (Russ, Daft, & Lengel, 1990). Brutus (2010) pointed out that since narrative comments are usually specifically and uniquely provided for individuals, they do surpass numeric ratings in their capacity to deliver a sense of individual focus to recipients. Therefore, performance ratings are not considered as personalized as written comments. Furthermore, in organizational communications, a "lean" medium such as standard report is more suitable for delivering routine messages addressed to large audiences, while a "rich" medium is best suited to transmitting nonroutine or individualized messages (Russ et al., 1990). Therefore, richness of narrative comments makes them more suitable for communicating feedback information, which is a nonroutine task in organizations (Daft & Lengel, 1986).

Qualitative data used to be deemed very expensive and difficult to process because the collection, coding and analysis consume considerable time and labor. However, thanks to the fast development of technology, it is no longer that hard and costly to collect, analyze and report narrative comments (Geer, 1991; Jackson & Trochim, 2002; Mossholder, Settoon, Harris, & Armenakis, 1995). As an example, computer-aided

textual data analysis (TDA) processes data through such activities as "listing, counting, or sorting words, phrases, and statements in the text, comparing words with a 'dictionary' of terms, adding reference information, identifying keywords, and making or coding key text segments" (Mossholder et al., p. 336). Also, as introduced by Jackson and Trochim (2002), concept mapping methodology blends the advantages of word-based and code-based analysis approaches while minimizing some of their weaknesses, and especially works well with those texts answering open-ended survey questions. This methodology analyzes content by creating units of analysis, sorting them into similar concept groups, generating cluster solution and labeling the clusters. These are just some examples of method that can be used to simplify qualitative data analysis.

In terms of how to measure narrative comment, it may be necessary to underline its special characteristics in performance appraisal. As a supporting part of performance appraisal, the comments may provide "rich description of respondent reality" (Jackson & Trochim, 2002, p. 307). From a linguistic perspective, descriptive text involves observed object, background and purpose (Lavinia, 2004). Written comments are produced to describe the evaluatee within the organizational context for the purpose of performance appraisal, which well distinguishes them from general type of text. Relja (2007) pointed out that the length of descriptions does not convey the complexity of different phenomena because it ignores the relationship between the components of a phenomenon. Rather, the complexity of the text depends on the content of the information that it conveys. In performance appraisal, comments are supposed to describe performance-related behaviors or results, which represent the most important information to be delivered to the feedback user or recipient. Therefore, when it comes to measure, the

content of comments in performance appraisal needs to be understood and treated differently compared to other types of text.

Intuitively, one may think that the amount of comments can be measured by the length of the text. However, as discussed above, the text length, such as number of words or sentences, would be inappropriate to gauge written comments produced in performance appraisal. Harman (2009) also questioned the value of assessing the length of comment. On one hand, lengthy comments may contain more information than short comments, but on the other hand, lengthy comments may simply just be repetitive. For this reason, comment length could generate a misleading interpretation. Comments tend to reflect performance-related behaviors or results and the meanings of the comments are expected to be pertinent to the organizational context (Jones & Kayongo, 2008; Smither & Walker, 2004). Therefore, it would be more appropriate to identify the amount of comments according to the meanings that pertain to the need of performance evaluation, i.e. the "narrative unit". Here, "narrative unit" specifically refers to the "meaningful and distinguishable performance-related theme found in narrative comments" (Brutus, 2010, p. 147). A narrative unit can be multiple sentences, one sentence, one phrase or even one single word, as long as its content contains one and only one performance-related theme.

Also, a deficiency has been noticed at the early stage of performance evaluation practices, that is, managers have tended to appraise subordinates by their personality traits, abilities or knowledge, rather than their actual work performance, or they simply have considered personality traits, abilities or knowledge as a part of work performance (Sims, Veres III, & Heninger, 1987). Whetten and Cameron (2005) stressed that management skills should not be identified by personality traits, but should be identified

by specific behaviors of individuals. The observable dimensions of effective work performance tend to be independent of and identifiable across individuals of different styles and characteristics. Smither and Walker (2004) also proposed that performance feedback is more effective when it focuses on work-related task or behavior than on recipient's personal traits or characteristics. Through a 3-year study of multi-source feedback for 100 managers, Miller, O'Farrell and DeMay (2006) confirmed that behavior-focused comments were most beneficial to improving employees performance and should be encouraged through training for evaluators.

The specificity of comments was labeled "comment scope" in Harman's (2009) study, that is, a general scope is broad in nature, and a specific scope focuses on particular issues. However, her codings of this variable were simply based on coders' subjective interpretations as "not at all specific", "somewhat specific" and "very specific", which appeared to be a tenuous measure. In contrast, in Smither and Walker's (2004) study of effects of comments on managers' performance, the specificity of comments was operationalized by the degree of task/behavior versus trait focus. In comparison, a measure by the degree of behavior versus trait is more appropriate as it focuses more on work-related dimensions. To put it another way, the more a comment focuses on behavior or task, the more specific it is considered to be.

As discussed above, composition of comments should not be overlooked in the evaluation process. This study will examine the influences of personality traits on production of comments. The theoretical framework will be introduced next.

Theoretical Framework

Big-Five Factors of Personality Traits

In behavioral science, personality has long been used as a fundamental individual difference, since it is relatively inherent, visible and distinguishable. Back to the late nineteenth century, scientists had recognized and classified in the dictionary up to one thousand words that describe personality traits of human beings and attempted to investigate the relationship between these personality terms and human behaviors (Goldberg, 1990). During the twentieth century, through continuous testing efforts by researchers, the lexicon of personality traits has been sharpened and refined to a better accepted taxonomy to satisfy the needs of behavior prediction, which is generally called the "Big-Five Factors". The validity of the model was well supported by employing different theoretical frameworks and instruments, testing across various cultures, adopting ratings from multiple sources and using a variety of samples, resulting in broad attention and increasing interests within the psychological field (Barrick & Mount, 1991; Digman, 1990). The "Big-Five factors of personality traits" is considered to be the most systematic and influential model of human personality in the psychology domain (Barrick & Mount, 1991; Digman, 1990; Goldberg, 1990, 1993; John, 1989; McCrae & Costa, 1989) and it has been widely accepted by researchers and frequently employed in studies related to human personality traits. It provides a grounding framework for "testing hypotheses relating individual differences in personality […] especially in the subfields of personnel selection, performance appraisal, and training and development" (Barrick & Mount, 1991, p. 23).

The Big-Five factors of personality traits model are usually defined as Openness to Experience, Conscientiousness, Extraversion, Agreeableness and Neuroticism (or Emotional Stability) (Costa & McCrae, 1992). The first dimension, Openness to Experience, is usually associated with appreciation for variety of experience, being open-minded to new ideas, full of curiosity and imagination, and artistically sensitive. Opposite traits would be described as being conservative, traditional and close-minded. The second dimension, Conscientiousness, on one hand, refers to traits pertaining to dependability, such as being organized, responsible and thoughtful; on the other hand, it reflects volitional characteristics like being hardworking, adherent, achievement-oriented, which is more associated with willingness for accomplishment (Barrick & Mount, 1991). The third dimension, Extraversion, is characterized by active, energetic and sociable personalities. Extraverted people tend to be expressive, enthusiastic and assertive. They enjoy the company of other people, like to be involved in social activities and draw attention to themselves. In contrast, introverted people tend to be quiet and low-key in social occasions. However, that does not necessarily mean they are shy. They may just simply be less socially active because they feel more comfortable on their own and enjoy tranquility more. As to the fourth dimension, Agreeableness, it is usually associated with a high degree of social conformity and trust in human nature. Such people tend to be cooperative, caring and forgiving to others. Social harmony is an important value to them. Therefore, they usually behave in a friendly and tolerant way, show concerns for others' feeling and are willing to compromise their opinions or behaviors with others. The last dimension, Neuroticism, refers to a tendency for experiencing negative emotions such as

anger, depression or anxiety. They are more sensitive and vulnerable to stress or adversity and behave more emotionally unstable.

Extensive use of the Big-Five factors of personality traits as an important theoretical basis to explore their connection with human behaviors can be found in the past two decades. Within the field of business administration alone, there exists an extensive body of literature exploring the relationship between the Big-Five factors and all kinds of behaviors in organizational settings, such as job performance, organizational citizenship behavior (OCB), personnel selection as well as team performance. To give an example, in a study on the impact of personality traits on job applicants' success, Caldwell and Burger (1998) conducted an interesting research among graduating college students who were searching jobs. They found that those students who scored higher in Openness to Experience, Extraversion, and Conscientiousness tended to make better use of social sources to prepare for their interviews, i.e. obtaining background information of the company or the job from friends, relatives or other kinds of social connections. Furthermore, Conscientiousness was found to be positively related to the use of non-social preparation, i.e. obtaining information of the company from written material. Students who better used social sources for preparation for initial interviews were more likely to have follow-up interview opportunities and receive job offers, indicating that there was a relationship between personality and applicant's success through their preparation before the interviewing process. Moreover, the interviewers' perceptions of the interviewees' personality had an impact on the job search success as well.

Big-Five Factor Model and Performance Appraisal

In organizational settings, individual differences do not only influence employees' job performance but also affect their tendencies or behaviors as evaluators (Tziner, Murphy, & Cleveland, 2005). It is not rare to find the Big-Five factor model applied in performance appraisal research. According to the previous results from some researchers, the personality attributes of evaluators can influence the outcome of their evaluations (e.g. Kane, Bernardin, Villanova, & Peyrefitte, 1995; Bernardin, Villanova, & Cooke, 2000; Tziner, Murphy, & Cleveland, 2002; Yun, Donahue, Dudley, & McFarland, 2005).

Bernardin et al. (2000) conducted a study among more than one hundred undergraduate students, aiming to find out the impact of raters' personal characteristics on their rating elevation. According to the Big-Five factor model, conscientious people are described as diligent, advertent and disciplined, usually set high performance standards for themselves and even others, and agreeable people are deemed tender-hearted, considerate and forgiving, value social harmony and tend to avoid conflicts. Therefore, Bernardin et al. predicted that those people who were higher on Agreeableness factor and lower on Conscientiousness factor would tend to show relatively more leniency in their appraisals, leading to elevated and inaccurate evaluations. Results were obtained from the students' evaluation for their peers and the speculations were supported. As an extension of Bernardin and his colleagues' study, Yun et al. (2005) did further research on the relationship between traits of Agreeableness and Conscientiousness and rating behavior, by taking social context (i.e. whether or not evaluation is followed by face-to-face feedback interaction) and rating format (i.e. graphic rating scale vs. behavioral checklist) into account as potential moderators. These

authors also found a positive relationship between Agreeableness levels of raters and the rating levels they provided, while both social context and rating format did have moderating effects. For instance, when one high on Agreeableness expected a face-to-face conversation with the ratee about the feedback he or she provided in the evaluation, the person was more likely to elevate his or her ratings. This corresponded to a tendency of being concerned about others' feelings and avoiding interpersonal confrontation characterizing the trait of Agreeableness.

Among the Big-Five factors of personality traits, Agreeableness and Conscientiousness have been studied most frequently in research on evaluation behaviors. Agreeableness is commonly considered to be positively associated with rating leniency (e.g. Bernardin et al., 2000; Kane et al., 1995; Kuang, 2005; Spence & Keeping, 2010; Yun et al., 2005), and Conscientiousness is usually expected to be negatively associated with rating leniency (e.g. Bernardin et al., 2000; Kuang, 2005; Spence & Keeping, 2010; Yun et al., 2005) or beliefs about the uses of performance appraisal (e.g. Tziner et al., 2002).

It is also believed that personality traits play a role in performance appraisals because they are likely to affect one's motivation to write comments about others (Brutus, 2010). Using the Big-Five factor model as the theoretical framework, this paper will examine the relationship between certain personality traits of the performance appraisal users and their production of narrative comments.

Purposes of Research

This study intends to examine the relationship between certain personality traits and the users' production of narrative comments in performance appraisals. To be more specific, the purpose is to understand how the traits of Extraversion, Agreeableness and Conscientiousness of a person would contribute to his or her production of narrative comments, including the amount (quantity), the specificity (quality) and the positivity of the comments. Individual personality tends to be inherent and unchangeable, which means it is unlikely to alter when one's organizational context or position switches. This study aims to offer some perspective for performance qualitative feedback recipients as well as human resource professionals in organizations to help them understand the influences of personality on comments. Last, by discussing the implications of the findings, this paper calls for more extensive research effort to the use of qualitative data in performance appraisal.

Research Model and Hypotheses

Among the Big-Five factors, it is proposed that Conscientiousness, Extraversion and Agreeableness will influence the amount and the content of narrative comments. In the research model proposed, these three factors represent the independent variables; the dependent variables being amount, specificity and positivity of narrative comments. In the next section, I describe in more detail the rationale for the proposed hypotheses.

Conscientiousness and Writing Comments

Before writing comments, the evaluator needs to think about the following questions: To whom the comments are addressed (i.e. the audiences of the comments)? What are the comments about (i.e. the content of the comments)? What outcomes are the comments expected to generate (i.e. the purpose of the writing) (Grabe & Kaplan, 1996)? In performance evaluation, written comments are addressed to evaluation receivers such as co-workers, team members or subordinates, containing performance-related information. The main purposes are to serve administrative needs in organizations and support managerial decision-making process, and at the same time, to satisfy informational needs of both evaluation users (practitioners) and receivers (evaluatees). Evaluation information is not only useful for management but also for evaluatees, who expect emotional assurance about their work performance, if they think they did a good job, or reflect upon feedback that they receive, if they think they need further development.

Comments as product and writing as process should be seen as "complementary perspectives on the same phenomenon" (Grabe & Kaplan, 1996, p. 82). For a text producer, he or she needs to keep linguistic information organized to make textual choices in his or her writing (Lavinia, 2004). Berge (1997) found that students worked harder mentally and learned more when they needed to put their thoughts in writing. Similarly, compared with ratings, writing narrative comments about others' performance requires more cognitive effort (Brutus, 2010). The writer needs to retrieve from memory specific and precise information about the evaluatee's performance, organize the information and transform it into words of evaluation, or even provide additional suggestions for improvement. In another word, writing comments is made up of two key

components: cognitive effort for information recollection or organization and cognitive effort for text production.

In general, conscientious people usually have a high sense of responsibility, like to follow schedules and strive for achievement, and they are usually self-disciplined and hard working (Costa & McCrae, 1992). The trait of Conscientiousness is often used to predict performance level. For instance, it has been frequently linked to and is believed to be the most prominent Big-Five predictor of academic performance (De Raad & Schouwenburg, 1996; Poropat, 2009). Conscientious students tend to demonstrate higher academic achievement compared to those who are less conscientious, resulting from the amount of time and effort they put in school work (Bidjerano & Dai, 2007; De Raad & Schouwenburg, 1996; Noftle & Robins, 2007; Trautwein, Ludtke, Roberts, Schnyder, & Niggli, 2009). With a high sense of responsibility as well as self-discipline, conscientious individuals are driven by duties and fulfill obligations voluntarily for interests of others. On the other hand, they will also choose to perform well in achievement tasks out of self-interest. As a result, they do not shirk from duties and input persistent effort, no matter with or without incentives (Fong & Tosi, 2007). In Bidjerano and Dai's study about the relationship between personality factors and self-regulated learning skills, they found that Conscientiousness was significantly related to students' efforts to optimize their use of time as well as mental concentration while studying. Seen from above, it is not difficult to find evidence for the connection between Conscientiousness and willingness to put effort. Therefore, when it comes to writing comments, it is expected that conscientious individuals would expend more effort, i.e. spend more time thinking and writing, to retrieve accurate information about the evaluatee and produce quality comments to fulfill

the duty as an evaluator.

Also, the trait of Conscientiousness is often associated with studiousness, organization and deliberation, which means that conscientious individuals are attentive to details and well organized (Costa & McCrae, 1992). Previous research by Digman and Inouye (1986) mentioned that, to some degree, Conscientiousness could be an indicator of one's attentive capacity. Therefore, it is believed that, on one hand, a conscientious person would pay closer attention to details of others' performance in workplace and be better able to make a mental record of performance-related incidents. He or she will remember more clearly and recollect more relevant information before writing comments. On the other hand, as a writer, he or she would try hard to organize his or her thoughts well and translate them into comments.

In summary, it is hypothesized that conscientiousness leads to a greater amount and more specific written comments. Partial evidence was provided by a very recent study conducted by Harman (2009) about open-ended survey comments on a training evaluation. In this study, it was found that conscientious respondents provided more specific comments.

Hypothesis 1: Conscientiousness is positively related to the amount of narrative comments produced in a performance appraisal context.

Hypothesis 2: Conscientiousness is positively related to the specificity of narrative comments produced in a performance appraisal context.

Extraversion and Writing Comments

Extroverted individuals tend to be outgoing and socially active. They like to communicate with others and express their ideas or opinions, and therefore usually engage in social contacts more easily (Costa & McCrae, 1992). While engaged in speeches or conversations, compared to introverts, extraverts show a tendency of using a greater number of words, i.e. talking more and for a longer period of time (Carment, Miles & Cervin, 1965; Gifford & Hine, 1994). Mehl, Gosling and Pennebaker (2006) used Electronically Activated Recorder (EAR) to capture 96 students' interactions and language use for two days. Evident expression of Extraversion was found in their daily activities as the recording revealed that extraverted students participated in more conversations and spoke more words over the period, relative to those introverts.

Nevertheless, some research has shown that proactive communication tendency does not only influence oral expressions but also written expressions in text-based communications. Especially with the emergence of telecommunication and computer-mediated communication (CMC) (Stefanone & Jang, 2007) tools such as Short Message Service (SMS), email or MSN, communications are no longer limited to traditional face-to-face or telephone conversations but are more often realized through these new media.

Extraversion has been recently studied in a couple of interesting linguistic studies. One example is Oberlander and Gill's (2006) study about language use. They were interested in finding out the effects of individual differences on people's specific linguistics choices in their email writings. One hundred and five participants in the study each wrote two email texts according to the assigned topics. The researchers analyzed the syntactic and wording tendencies in the collected email texts, with predictions involving

distinctions of positivity, sociability, complexity and implicitness between individuals. Implicitness was defined as informality and "a preference for pronouns, adverbs, and verbs, as opposed to nouns, adjectives, and prepositions" (p. 242). They expected higher positivity, sociability and implicitness in the texts written by extraverts than those written by introverts. The results indicated that, first, extroverted individuals tended to use more adjectives, and had more collocations (i.e. words or terms that often go together) associated with certainty such as "be really" and "be so" in their writings. Secondly, introverted individuals used adjectives less frequently and showed more tentativity (uncertainty). The characteristics of extraverts' writing habits exhibited in this study were specifically related to grammar and syntax; this provides an interesting perspective on the connection between Extraversion and writing behaviors.

Oberlander and Gill (2006) studied personality data obtained from self-report questionnaires instead of reports by third-party readers of the writings. As a counterpart, Vignovic and Thompson (2010) studied email recipients' perceptions of senders' characteristics in face of linguistic errors, i.e. grammatical and spelling mistakes, and etiquette violations, i.e. impolite messages, that occurred in emails. It was found that both linguistic errors and etiquette violations influenced the recipient's perception of sender's levels of extraversion and conscientiousness.

Stefanone and Jang (2007) examined the effects of personality on the use of blogs, a kind of interactive text-based posts that express personal feelings and thoughts on Internet. Similar to personal diaries, blogs provide a virtual space for bloggers to write about their daily activities or express personal opinions. The difference is that blogs can be shared with other bloggers and receive comments from them. As a result of the

interactive nature of blogs, bloggers often have an idea who are the audiences and expect to see comments about their own writings from the viewers. Stefanone and Jang's study found that people high in extraversion and self-disclosure (a tendency to disclose personal information to others) were more motivated and better able to capitalize this online writing tool for maintaining their social networks with friends or family, which satisfied their psychological needs of "relationship development and maintenance".

From all the research above we can see that extrovert's distinctive personality is not only closely related to their oral discourses but also to their writing behaviors.

As to performance appraisal, research about the relationship between Extraversion and writing comments seems to be scarce. Narrative comments stand for "the need to communicate one's evaluation in writing" (Brutus, 2010, p. 148). Evaluators are clearly aware of to whom they are going to "talk" as well as the communicating purpose through written comments. On one hand, the most evident characteristic of an extroverted person is being active in expression, no matter through conversations in person or texts on paper. It is predicted that, during the process of expressing one's opinions about others' performance, an extravert would give more of his or her thoughts about others in the evaluation compared to an introvert, and consequently, write more and more specific comments. On the other hand, Extraversion was proved to be a predictor of memory. Memory has been defined in terms of short-term and long-term memory. Short-term memory refers to memory retention after a period shorter than 10 minutes, and long-term memory is memory retrieval after a longer period of time (Lieberman, 2000). It has been suggested that individuals high in Extraversion have advantages in both long-term (Baker & Bichsel, 2006) and short-term memory (Pearman, 2009) that were demonstrated in

cognitive tasks such as number, name and image retrieval and recognition. Gray and Braver (2002) also suggested that an attribute similar to Extraversion, which is called behavioral approach sensitivity, was able to predict a better performance of working memory, i.e. a mental capacity for active maintenance and manipulation of information. It can be inferred that extraverts would be better able to retain a mental record of performance cues in mind. At the time of writing comments, they can remember more clearly, retrieve more relevant information about others' performance and use the information accordingly.

All the evidence above suggests that, relative to introverts, extraverted people are likely to write more comments with more details and therefore leads to the hypotheses as follows:

Hypothesis 3: Extraversion is positively related to the amount of narrative comments produced in a performance appraisal context.

Hypothesis 4: Extraversion is positively related to the specificity of narrative comments produced in a performance appraisal context.

Agreeableness and Writing Comments

People who are high on agreeableness tend to value harmony in social relationship and would compromise their interest, ideas or opinions with others. They are also sympathetic to others and don't like to be critical (Costa & McCrae, 1992). The influences of Agreeableness trait on a person's writing behavior can usually be demonstrated in the person's willingness and tone to express his or her opinions. First,

people high in Agreeableness are less active in giving opinions to others and more tolerant with status quo. Second, Agreeableness is closely related to positive emotional expressivity including expression by language (Gross & John, 1997). Mehl et al. (2006) recorded a higher amount of arguing among less agreeable female students in their tracking study, which evidenced the negative oral expression tendency among disagreeable individuals. Also, they found a higher frequency of using first-person singular pronouns such as "I, me, my" for individuals high on agreeableness, which symbolize a language form that expresses consensus (Pennebaker, Mehl, & Niederhoffer, 2003).

LePine and Van Dyne (2001) considered expression of one's voice as a form of contextual performance and a kind of change-oriented behavior in organizations, since it usually aims to express dissatisfaction, initiate improvement or provide constructive suggestions. These authors believed that agreeable individuals would not sacrifice interpersonal relationships for facilitating changes and would rather more likely accept the status quo. Their laboratory study did prove the negative relationship between the degree of Agreeableness and individual's voice behavior, i.e. a lower tendency to give different "voices" of agreeable people.

Although as stated above, an agreeable individual is relatively reluctant to express opposing opinions to others, it does not necessarily mean that the person will not give any comments in performance evaluation. He or she may just express opinions in a manner that will not upset others' feelings. Therefore, it cannot be predicted that there is a relationship between Agreeableness and amount or specificity of comments, and this trait is more likely to be related to positive expressivity. That is to say, when writing

comments, it is believed that agreeable individuals are likely to express their opinions about others in a more positive tone or with more positive language, rather to criticize in a harsh tone, because of their inherent tender-mindedness and their avoidance of social conflicts. Interestingly, in Harman's (2009) study about survey comments, her prediction of a positive relationship between Agreeableness and providing positive-toned comments was not supported. Instead, it was the respondents' satisfaction level that was found positively related to the tendency to provide comments in positive tones. Nevertheless, another attempt will be made in this study to test such a hypothesis.

Hypothesis 5: Agreeableness is positively related to the positivity of narrative comments produced in a performance appraisal context.

Group Size and Writing Comments

It is predicted that the relationship between personality traits and production of comments will be moderated by the size of the group that the person is in; that is the smaller the group, the stronger the relationship is likely to be. To put it another way, in smaller groups, Conscientiousness or Extraversion will have a stronger effect on team members' production of comments. In organizational behavior studies, team size has been linked to challenges of communication and coordination. A larger team is more likely to be confronted with difficulties and higher costs of communication and coordination because of more complex linkages among the members (LePine, Piccolo, Jackson, Mathieu, & Saul, 2008). In Bradner, Mark and Hertel's (2005) research about team size in a large multinational organization, results indicated better knowledge of team members and more active participation occurred among members of smaller teams.

As team size increases, team members will have fewer opportunities to participate in team activities such as meetings, and interactions within the team become more complex as well, which consequently results in less interpersonal communications and lower awareness with one another.

During the course of the group project, team members have opportunities to interact with one another through communications such as group meetings, email contacts and casual conversations. When one writes comments about a certain team member, he or she needs to retrieve information from those interactive activities in his or her memory before identifying the performance of that team member and putting the retrieved information into words. In a relatively large team, the disadvantage of more complex interactions is likely to make that writing process more difficult. Therefore, the more group members a person needs to write comments about, the more mental and physical effort he or she needs to make and the more time to take. Consequently, the influence of a large group is going to counteract that of personality such as Conscientiousness or Extraversion on the quantity and quality of the comments. For instance, a team of 6 members have more complicated interactions during the group project compared to another one of only 3 members, resulting in more effort and time that are demanded for writing comments. Due to the large group size, Conscientiousness or Extraversion is likely to underplay in the production of comments, and thus the team members would probably write fewer and less specific comments than they would have.

Hypothesis 6a: Group size moderates the relationship between Conscientiousness and the amount of narrative comments. The relationship is stronger for smaller groups.

Hypothesis 6b: Group size moderates the relationship between Extraversion and the amount of narrative comments. The relationship is stronger for smaller groups.

Hypothesis 7a: Group size moderates the relationship between Conscientiousness and the specificity of narrative comments. The relationship is stronger for smaller groups.

Hypothesis 7b: Group size moderates the relationship between Extraversion and the specificity of narrative comments. The relationship is stronger for smaller groups.

The research model and the hypotheses are illustrated below respectively as Figure 1, 2 and 3.

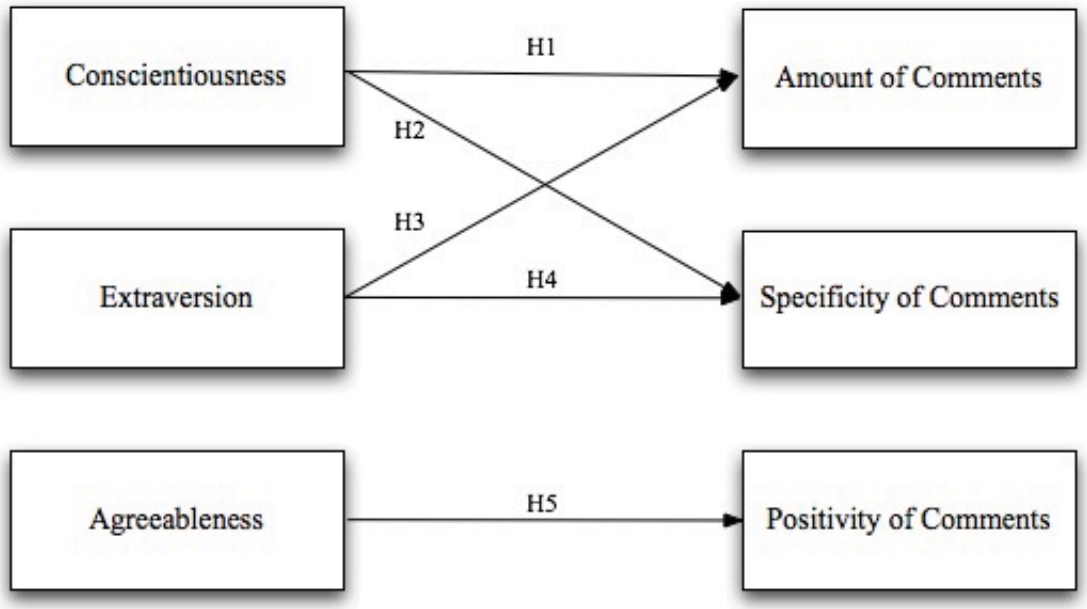

Figure 1 Hypotheses 1, 2, 3, 4 and 5

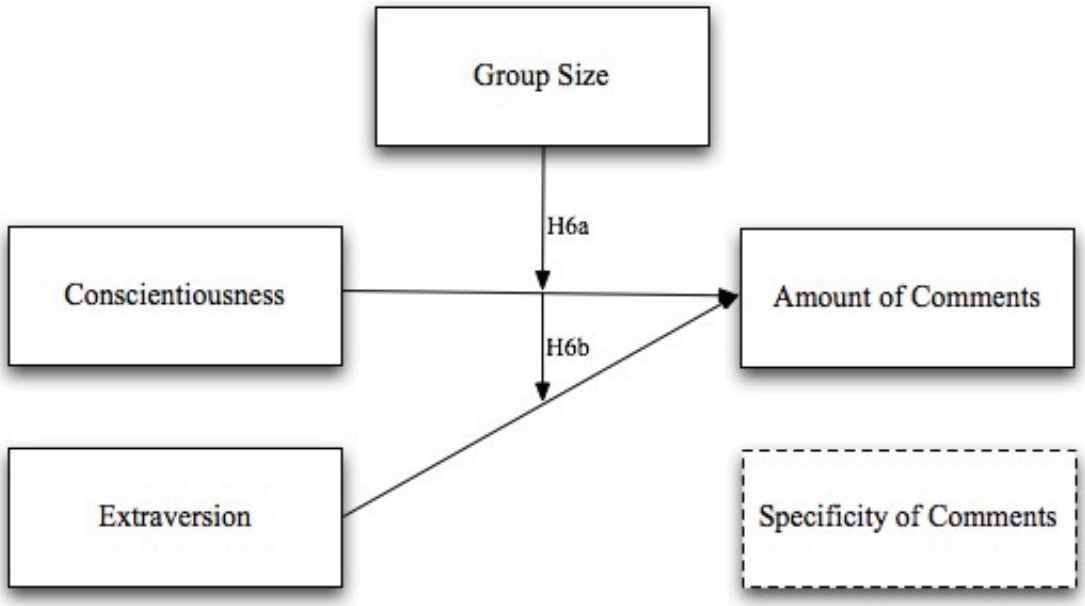

Figure 2 Hypotheses 6a and 6b

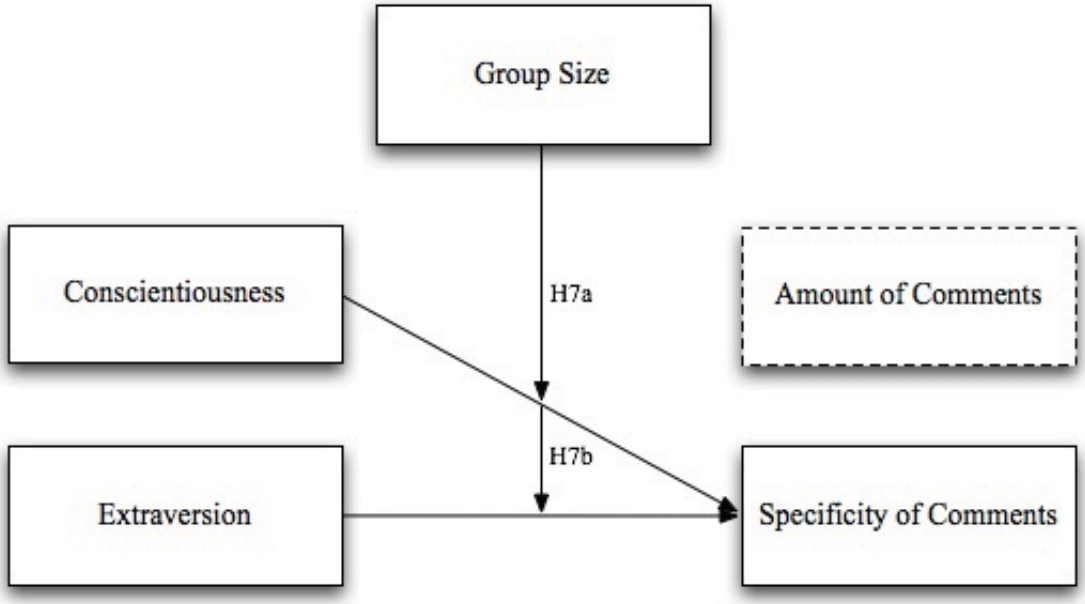

Figure 3 Hypotheses 7a and 7b

Research Methodology

Sample

The study was conducted at the John Molson School of Business of Concordia University in Montreal, Canada. The sample of the study was undergraduate students from the COMM 222 subject pool who used an electronic peer evaluation system for their group projects at the end of the term. They participated in the study for 0.5 course credit.

During the two semesters, a total of 110 students completed the web-based personality survey, however, some of them did not provide their student IDs, which means their personality scores could not be linked to their comments. After eliminating the missing data, the final sample consists of 87 undergraduate students, including 36 male (41% of the total sample) and 51 female (59% of the total sample) students. There are 31 students from Accounting, 15 from Administration, 4 from Economics, 14 from Finance, 2 from HR Management, 3 from MIS, 3 from International Business, 2 from Management, 8 from Marketing, 3 from Supply Chain Management, 1 from Psychology and 1 from Political Science. In addition, some students only completed the numerical ratings of the peer evaluation without providing any narrative comments. For these data, the amount of narrative units was coded as zero. When it came to the analyses of the specificity and the positivity of comments, those with missing comments were not included in the analyses and those participants were excluded from the sample. The sample size changed to 69, including 27 male and 42 female students.

Procedure

During the winter and summer terms of 2010, an electronic questionnaire using scales from the International Personality Item Pool (IPIP) (http://ipip.ori.org), which measures the levels of Conscientiousness, Extraversion and Agreeableness of the participants, was released on the HPR system of John Molson School of Business, where COMM 222 students can sign up for online surveys. The participants were informed of the purpose of this study and their rights as a research participant before they agreed to participate. When they completed the survey, they were also asked to provide their identities, i.e. student IDs. Data of the web-based survey was downloaded by the thesis author at the end of each semester. Subsequently, after the deadline of submitting electronic peer evaluation for group project, all the narrative comments given by the participants in the evaluation system were exported by the thesis supervisor. The final data delivered to the author for analysis, including personality traits and narrative comments, maintained anonymous and were linked up with each other by the participants' student IDs. Group size, major and gender of the participants were also obtained from the questionnaires.

To analyze the amount, specificity and positivity of the narrative comments, all the data were coded by the researcher and another graduate student independently. The coding protocol is attached as Appendix E. The other coder received about three hours of coding training from the researcher prior to independent coding.

Measure

Independent Variables

As mentioned above, the measures of the three independent variables, Conscientiousness, Extraversion and Agreeableness, are obtained from the IPIP.

There are 10-item and 20-item Big-Five scales available in the IPIP. Although the internal consistency of 20-item scale appears to be slightly higher than that of 10-item scale, 10-item scale is chosen to avoid a lengthy questionnaire. That is to say, each variable was measured by 10 items in a five-point likert scale (see the questionnaire as Appendix D). According to IPIP, the internal consistency reliabilities are .81 for Conscientiousness, .86 for Extraversion and .77 for Agreeableness (http://ipip.ori.org). Retest of reliability showed that the internal consistency reliabilities in this study were .77 for Conscientiousness, .87 for Extraversion and.71 for Agreeableness.

In every 10 items that measure each variable, there are 5 positively keyed (+ keyed) items and 5 negatively keyed (– keyed) items. For example, for measuring "Conscientiousness", the item "I make plans and stick to them" is positively keyed, and the item "I find it difficult to get down to work" is negatively keyed; another example would be that, for measuring "Extraversion", the item "I am skilled in handling social situations" is positively keyed, and the item "I don't like to draw attention to myself" is negatively keyed. The items are designed in this way in order to increase the accuracy of the measurement. Since + keyed and – keyed items represent opposite behaviors or personality traits of a person, while scoring the scales, the scores assigned to + keyed items are reverse to those assigned to – keyed items accordingly. To be more specific, in this study, for + keyed items, the response "strongly agree" was assigned a value of 5,

and the response “strongly disagree” was assigned a value of 1; on the contrary, for –keyed items, the response “strongly agree” was assigned a value of 1, and the response “strongly disagree” was assigned a value of 5.

The results showed that the score of Conscientiousness ranged from 2.30 to 5.00 (N = 87, M = 3.61, SD = 0.51), the score of Extraversion ranged from 2.40 to 5.00 (N = 87, M = 3.55, SD = 0.60), and that of Agreeableness ranged from 1.70 to 4.80 (N = 87, M = 3.71, SD = 0.43).

Dependent Variables

Amount of comments

The amount of comments was measured by the average number of “narrative units” in the comments. As mentioned earlier, narrative unit is conceptualized as meaningful and distinguishable performance-related theme (Brutus, 2010). To be more specific, the amount of comments was calculated by dividing the total number of “narrative units” in all the comments written by a student by the number of his or her group members.

Some comments contained many words or long sentences but only expressed meaning that was about a single theme which was related to the evaluatee’s performance such as “timeliness” or “cooperation”. That was rated as one narrative unit. Here is an example: “<u>Your sections could have been completed by assigned deadlines, for other team members may not like waiting until the very last minute</u>”; this statement is concerned with the the theme of “timeliness”. Another example “He always contributes new ideas in the project and always delivers his part of work in time” encompasses only

one sentence with two performance-related themes: “idea contribution” and “practical contribution”. Thus, this sentence was coded as containing two narrative units. Some comments did not touch any specific performance-related theme but only gave a vague evaluation such as “He is a good team member to work with” or “I am satisfied with her performance”. If that was the only comment the person writes, it was rated as one narrative unit, because after all it was a comment, although not a good one. If other comments followed as a further explanation of that point, i.e. why “he is a good team member”, in that case, this general evaluation was ignored and the remaining comments was rated according to the number of performance-related themes. For example, “Liyan was a good member to have in our group because he always had constructive ideas and wasn't afraid to point out any flaws.” The evaluation of “Liyan was a good member” was followed by explanations regarding “idea contribution” and “straightforward suggestion” : this was coded as two narrative units. Some more examples are given in Appendix E (Coding Protocol).

After rating the number of narrative units in each comment, the sum of narrative units in all the comments written by that student was added up and then was divided by the group size to obtain the average amount of comments given by each student (for per team member). The average amount ranged from 0.00 to 7.90 narrative units (N = 87, M = 1.65, SD = 1.26). After both the researcher and the coder finished coding, the interrater reliability was calculated as the intraclass correlation coefficient (ICC), which is the reliability of mean ratings from more than one coder (LeBreton & Senter, 2008). The interrater reliability for amount of comments was .98.

Specificity of comments

Specificity of comments was rated by the degree of behavior versus trait, which has been mentioned earlier. To be more specific, a five-point likert scale that was borrowed from Smither and Walker's (2004) study was used in this study, in which: 1 = comment that focuses entirely on personal traits or characteristics, 2 = comment that focuses mostly on personal traits or characteristics, but to a lesser extent also refers to specific behaviors or tasks, 3 = comment that focuses equally on personal traits or characteristics and specific behaviors or tasks, 4 = comment that focuses mostly on specific behaviors or tasks but to a lesser extent also refers to personal traits or characteristics, 5 = comment that focuses entirely on specific behaviors or tasks.

To give an example, "Josie really took the time to edit the paper for the final copy. She's a little quiet though, but still contributed to group ideas." This comment described Josie's behaviors of editing paper and idea contribution as well as her quiet characteristics. It was considered focusing more on behaviors than on personality traits and was rated 4. To give another example, "Did absolutely nothing, barely responded to any emails, handed her parts in extremely late and then they had to be rewritten. Made no effort taking any initiative or participating in the meetings. Never showed up to class and was difficult to get in touch with." The descriptions in this comment all focused on behaviors such as writing contribution and meeting participation and therefore it was rated 5. In some cases, there were some comments which were so general that focus on neither personal traits nor specific behaviors, such as "I enjoy working with him". This kind of comment was rated 3.

After rating the specificity of each comment, the ratings of all the comments written by that student was added up and then the sum was divided by the group size to obtain the average specificity of comments given by each student (for per team member). The average specificity ranged from 0.80 to 5.00 (N = 69, M = 2.72, SD = 1.16). The interrater reliability for specificity of comments was .97.

Positivity of comments

The Positivity of comments was rated by coders subjectively according to the degree of positive tone, no matter the comment was related to performance (behavior/task), personal trait, or any other dimensions. A five-point likert scale was used, in which: 1 = very negative, 2 = slightly negative, 3 = neutral, 4 = slightly positive, 5 = very positive.

To give an example: "You try hard and respect your peers. You contribute well to the thoughts but i believe that you must learn to better grasp certain subjects before you volunteer a large task like writing a body paragraph." This comment first praised the evaluatee's hard-working and respectful traits and idea contribution, and then provided a suggestion for improvement. It was rated 4 – slightly positive. Another example is "This was a group project and I feel like you were not very present during the whole process. You did not even show up the day of our presentation and I did not take that very well. Although you were cooperative, you did not really participate to the task at hand and were therefore not all that helpful. It would also be nice for you to should up on time." This comment criticized the evaluatee's absences and insufficient participation during the project process, with only one positive evaluation as "cooperative". Overall, the comment

was in negative tone. It was rated 2 – slightly negative. More examples are given in Appendix E (Coding Protocol).

After rating the positivity of each comment, the ratings of all the comments written by that student were added up and then the sum was divided by the group size to obtain the average positivity of comments given by each student (for per team member). The average positivity ranged from 0.60 to 5.00 (N = 69, M = 4.28, SD = 0.95). The interrater reliability for positivity of comments was .97.

Moderator

The group size was measured by the number of peers of the student (i.e. the evaluator) in his or her group, that is, how many people the student needs to write comments about for the group project. It ranged from 2.00 to 6.00 (N = 87, M = 3.74, SD = 0.84).

Results

Correlations among the key variables are presented in Table 1. The mean value of the amount of comments was 2.08, which means that in average the number of narrative units in participants' comments was relatively small. The mean value of the specificity of comments was 2.72, i.e. an average below medium degree of behavior versus trait. To be more specific, students provided comments that focused on behaviors slightly more than personal traits. The positivity of comments was high (*M* = 4.28). In other words, most comments tended to be in a positive tone. Furthermore, the average levels of participants'

Conscientiousness, Extraversion and Agreeableness were all above the mid-point (respectively 3.60, 3.56 and 3.77). There is some other research which used the same personality instrument with samples of students or employees (e.g. Ang, Dyne, & Koh, 2006; LePine & Van Dyne, 2001; Thoms et al., 1996; Tziner et al., 2002; Yun et al., 2005). According to the results in those studies, the mean values of Conscientiousness ranged from 3.33 to 4.06, those of Extraversion ranged from 3.17 to 3.50, and those of Agreeableness ranged from 3.29 to 3.57. In comparison, the levels of Extraversion and Agreeableness of the participants in this study appeared to be slightly higher. As shown in Table 1, those students who reported a higher level of Conscientiousness also declared themselves to be more extraverted ($r = .34$, $p < .01$) as well as more agreeable ($r = .43$, $p < .01$), and those who reported a higher level of Extraversion also declared themselves to be more agreeable ($r = .26$, $p < .05$). Additionally, those students who wrote more comments also wrote them with more focus on behaviors ($r = .43$, $p < .01$).

Hierarchical regression analyses were conducted to examine the hypotheses with gender entered as a controlled variable. In these analyses, all variables were centered (i.e., the mean of each variable was subtracted from each score) to reduce multicollinearity. For hypotheses 1, 2, 3, 4, 6a, 6b, 7a and 7b, on the first step, gender was entered as a controlled variable. On the second step, Conscientiousness or Extraversion was entered as the independent variable, and group size was entered as the moderator. On the third step, the product term of the independent variable and group size was entered to examine two-way interactions between them. For hypothesis 5, gender was also entered as a controlled variable in the regression.

Summaries of the regression results are presented respectively in Table 2, 3, 4, 5

and 6. First, for hypothesis 1, result indicates that Conscientiousness was not related to the amount of comments (N = 87, B = .12; t = .43; *ns*). Nor was it related to the specificity of comments (N = 69, B = .16; t = .59; *ns*). Hypothesis 2 was not supported, either. Secondly, not as expected, the trait of Extraversion was not related to either the amount of comments (N = 87, B = .20; t = .84; *ns*) or the specificity of comments (N = 69, B = .24; t = 1.08; *ns*). Thirdly, as opposed to hypothesis 5, it was found that Agreeableness was not related to the positivity of comments (N = 69, B = .24; t = .92; *ns*). Finally, according to the findings, none of the two-way interactions were significant. Group size had no effects on either the relationship between Conscientiousness and the amount of comments (N = 87, B = .09; t = .26; *ns*) or that between Conscientiousness and the specificity of comments (N = 69, B = .03; t = .07; *ns*). Group size also did not influence the main effect of Extraversion on the amount of comments (N = 87, B = .23; t = .94; *ns*) or that of Extraversion on the specificity of comments (N = 69, B = .19; t = .80; *ns*).

In sum, the hypotheses were not supported. That is to say, in this study, the level of Conscientiousness and Extraversion could not predict the amount or specificity of comments, and the level of Agreeableness could not predict the positivity of comments, either. Moreover, group size did not moderate the relationship between students' personality traits and their production of comments. However, it is worth mentioning that gender was found related to the specificity of comments (N = 69, B = .58; t = 2.06; $p < .05$; $\Delta R^2 = .06$).

Discussion

Quantitative data, i.e. numeric ratings, has prevailed in performance appraisal practices for its objectivity and convenience as well as its potential capability to satisfy the social interests of the users. Nowadays "rating" has almost become the equivalence of "evaluation" in a performance appraisal context. However, as discussed above, qualitative data, i.e. narrative comment, has its irreplaceable advantages of being descriptive, clarificatory, context-reflective, and behavior-informative. Compared to quantitative data, its richness in meaning communication and intensity of personal focus make it appeal to feedback recipients more and thus more suitable for delivering performance-related information. Since qualitative data is so important in performance appraisal, composition of comments should be stressed as an essential part of the evaluation process.

The purpose of this research was to examine the influences of personality traits on the comment construction and it was developed from the framework of Big-Five Factor Model. In particular, among the Big-Five factors, Conscientiousness, Extraversion and Agreeableness were believed to be most relevant to an individual's production of narrative comments. From a perspective of linguistics, similar to other writing behaviors, producing comments consists of two key components: cognitive effort for information recollection and organization, and cognitive effort for text composition. Specifically, the attributes of comments including amount, specificity and positivity were investigated. Here the amount of comments was measured by the number of "narrative unit", which was defined as the distinguishable and meaningful performance-related theme found in comments. As to the specificity, it was measured by the degree of focus on work-related

behaviors versus personality traits.

First, conscientious people tend to be hard working, attentive to details and well organized, and therefore would make more mental effort to retrieve memory and physical effort to write comments. Second, being active in expression relative to introverts, extraverts would be happier to communicate with the recipient by providing more opinions about their performance in the comments; also, Extraversion has been proved to be connected with memory so that it may contribute to a better mental maintenance of the recipient's performance details before writing comments. Third, agreeable individuals are likely to express their opinions in the comments in a more positive tone, due to their avoidance of social conflicts and positive emotional expressivity. Last, more prominent effects of these individual characteristics on the attributes of comments were expected in smaller groups. Since a large group usually involves more complex interactions, more effort would be demanded for writing comments, which is likely to counteract the influences of personality traits.

A strength of this study is the fact that the assessment of individual characteristics and the production of narrative comments were done at different points in time, which reduces common method variance. The use of a real evaluation process is believed to be another strength of this research effort. Nevertheless, the outcome of this study did not find support for these predictions. To summarize the final results, first, it was found that both traits of Conscientiousness and Extraversion did not appear to have significant influences on the amount or specificity of comments that students wrote in their peer evaluation, though they were positively correlated to the amount or specificity of comments. Secondly, agreeable students did not write their comments in a more positive

tone than those less agreeable ones. Third, group size did not seem to moderate the relationship between Conscientiousness/Extraversion and the amount/specificity of comments. Expected stronger effects were not found in smaller groups. In addition, gender was found related to the specificity of comments.

Several factors may have possibly caused the unsupported results, and they will be discussed as follows.

Social Context of Peer Evaluation

A major factor leading to the unsupported results could be the social context of peer evaluation. First, the peer evaluation in this study was conducted online. Such a web-based system made the evaluation process easier for students to complete as well as more efficient for administrators to manage, meanwhile it also allowed an opportunity for the evaluatees to view the comments about them. According to the result, the average group size of the sample was 3.71, which was not very large. Since linkages are relatively simple in a small group, members may easily be able to identify who wrote the comments. Especially when a student knows that other team members would have the opportunity to read what he or she writes, it is likely that even a disagreeable student would write positive comments in order to avoid social conflicts. Moreover, many undergraduate students attend a number of courses together during their undergraduate studies. For that reason, when a student expects continuous meetings or contacts with the current team members in the future, he or she may have more concern about social relationship maintenance. Under such circumstances, the influence of personality traits on production of comments could be overwhelmed by that of social needs. That is to say, the positivity

of comments may have been elevated. Secondly, students tend to be stressed about their academic work, particularly such as a group project that demands a lot of time and energy. At the end of each semester, the peer evaluation deadline is usually scheduled before the final exams, when the students need to focus on studying. The accumulated pressure may have counteracted their motivation or willingness to spend a lot of time or effort thinking thoroughly about how to write comments. Therefore, the amount or specificity of comments may have been affected.

Sample Limitations

The relatively small sample size could be a drawback. Although there are hundreds of students who are actively registered with the HPR system, and the participants were promised to be rewarded 0.5 course credit for participating the study as well, the response rate was not very high. On one hand, the number of students who were involved in group projects during the study period (one fall term and one summer term) was probably not very large in the first place; on the other hand, even those students interested in the study may have had hesitation about providing their identification information, i.e. student ID, for the study. Even provided with the reassurance of personal information protection, people may still had concern on their privacy disclosure. In future research, a larger sample size will be needed to generate more convincing results. A power analysis was conducted to estimate the necessary sample size. To estimate the relationship between personality traits and production of comments, the effect size was determined to be small. According to Cohen (1992), a small effect size is defined as .02 for testing multiple correlations. For 3 independent variables, as used in this study, a sample of 547 subjects

is necessary to detect an effect with power equal to .80 and at a significance level at .05. Seen from the gap between the sample size used in this study and the estimated value from the power analysis, it would be necessary to expand the sample size in future research.

Also, as the study was conducted with undergraduate students, the sample had its limit in representing a wider range of individuals. Further, the study was conducted in an educational institution. Discrepant results may occur in other types of organizations due to contextual factors. In this study, students used the peer evaluation system to evaluate their teammates' performance in group project, however, some of them did not seem to make good use of it, which will be discussed below. One of the main reasons could be that they did not think that using the system would make any differences in their own academic performance. At a real workplace, which is very distinctive from an academic environment, performance appraisal is usually implemented for administrative or personnel decision-making purposes. In that case, it could be quite another story.

Quality of Peer Evaluation

Last but not least, it is worthwhile to pay attention to the quality of peer evaluation data. Some students only wrote comments as brief as one sentence or even one word. Also, it was not rare to see comments that were simply about personal feelings such as "I enjoy working with her" or solely about personal traits such as "he was a smart guy". The mean values of the amount and the specificity of comments were respectively 2.08 and 2.72. The relatively low values indicated that many participants failed to provide comments that were well composed or pertinent to performance-related tasks or

behaviors regarding the group project. Although guiding descriptions had been posted on the front page of peer evaluation system to request students to provide comments about specific performance aspects, such as “cooperation”, “conceptual contribution”, “practical contribution” and “work ethic”, many final products of narrative comments were still below expectation.

On one hand, students’ intense academic workload may have led to the unsatisfactory quality of comments; on the other hand, the outcome could have resulted from insufficient training for evaluation process. To improve the quality of narrative comments in performance evaluation, it is believed that it would be necessary to emphasize the importance of evaluation training and guidance. Evaluators need to understand that narrative comments should avoid to be personality trait oriented or simply about personal feelings for the evaluatee, but focus on performance-related tasks or behaviors. It is not fair to measure one’s performance only by emotional feelings or individual characteristics. Comments that describe specific behaviors would be more informative, convincing and instructive.

Future Research

In this study, self-reports were used to rate the participants’ personality characteristics. Although this method is commonly employed in personality surveys (e.g. Ang, Dyne, & Koh, 2006; LePine & Van Dyne, 2001; Thoms et al., 1996; Tziner et al., 2002; Yun et al., 2005), it is suggested that future research try combining self-reports with third-party ratings. Although compared to the results from some other research that

used the same personality instrument, the average scores of Conscientiousness, Extraversion and Agreeableness of the participants in this study were not largely discrepant, the survey respondents still could have elevated the ratings more or less subconsciously. After all, few people would want to admit to be an irresponsible and lazy student. In the meantime, they may also think an introverted person are usually labeled as shy, boring and unpopular to others. Therefore, in order to optimize the accuracy, self-ratings combining with ratings by others may be considered in future research if possible.

In addition, although no significant influences of the personality traits on the features of comments were detected in this study, it was interesting that gender appeared to be related to the specificity of comments. The average specificity of comments provided by female students was 2.4 and that provided by male students was 2.9, which means that male students tended to write comments focusing slightly more on behaviors than female students. After discovering this result, I did some research for studies on gender differences demonstrated in communication style and found mixed opinions. Some studies indicated that females are more expressive than males (e.g. Indartono & Chen, 2010; Jolson & Corner, 1997; Rosen, Chang, Ervin, Carrier & Cheever, 2010). For instance, one of the noticeable personality traits that related to sales performance has been instrumental/expressive (I/E) characteristics, with expressiveness typically representing feminine traits (Jolson & Corner, 1997). Women are not only orally but also verbally expressive. For example, Rosen et al.'s study (2010) showed that young female adults sent a considerably greater amount of text messages in their daily life than male young adults did. While some research revealed that males intend to play active and leading roles in communication (e.g. Selfe & Meyer, 1991; Stewart, Shields, Monolescu,

& Taylor, 1999). Stewart et al. (1999) found that, when using an on-line communication tool, in each of the three comparison groups, men always sent more and longer messages than women. Research also showed that, among senior organizational executives, male leaders reported to be significantly more expressive than female leaders (Callahan, Hasler, & Tolson, 2005).

Although there exist competing opinions for gender differences in communication style, it is believed that, the difference found in this study actually stems from the focal point of males' and females' communication process. It is already known that women are more sensitive to details than men are when their brains receive and process information (Krugman, 1966; Meyers-Levy & Maheswaran, 1991; Meyers-Levy & Sternthal, 1991). Meyers-Levy and Maheswaran (1991) conducted an experiment to explore the information processing strategies adopted by both genders, and found that females showed a higher tendency and sensitivity toward message contents and thus were able to elaborate more details when they were retrieving and recognizing message cues. In contrast, males tended to adopt a strategy to think and respond in an abstract and integrated manner. If that is true, one may think that females should write more specific comments since they pay closer attention to details than males do. Nevertheless, more importantly, women's attention to details is more likely to be drawn to personal feelings or experiences as shown in their conversations or writings (Jo, 1990; Taylor-Nelms, 2003). For instance, self-referring pronouns such as "I" or "me" appear frequently in women's language. Also, they find connectedness in one's relationship to others. Within the organization, women focus on more personal issues rather than fitting themselves into the constitution (Schultz & Anderson, 1986). Jo (1990) examined several hundred essays

written by students and found that female students showed more concerns with interpersonal relationships in their writings. However, for men, they like the contents of their communication to be impersonal and tend to view issues in an abstract way (Schultz & Anderson, 1986; Tannen, 1990). In the workplace, pronouns like “we” or “they” is often used in their expressions, and they tend to see themselves connected to the organization. They also talk about their work in more abstract terms (Schultz & Anderson, 1986). As evidenced by Tannen’s (1994) study, men like to draw their conclusions by referring to impersonal, concrete and task-related information because they think that opinions established on personal experiences lack persuasion. Therefore, such a perspective may help explain why male students wrote comments focusing more on behaviors than female students in this study.

Previous literature that examined gender effects on performance appraisal usually focused on differences of rating level between men and women when they act as raters or ratees (e.g. Fletcher, 1999; Furnham & Stringfield, 2001; Scherer, Owen, & Brodzinski, 1991; Shore & Thornton, 1986). Future research could spend some effort in addressing gender effects on writing behaviors in performance appraisal, in terms of the amount and the content of comments.

By looking into personality traits and their effects on production of comments in performance appraisal, this study aimed to provide a new perspective on the relationship between writers and comments. Despite the unsupported results, it is believed that this study still has its value for its attempt to explore a new field about qualitative data in performance appraisal. Suppose that characteristics such as Conscientiousness, Extraversion or Agreeableness did have influences on the writer’s production of

comments, it would help us understand how some specific differences in narrative comments are caused by individual personality distinctions. Due to some factors such as social context of peer evaluation and sample limitations, evidence was not found to support the predictions. In future research, it may be necessary to try to minimize the effects of such dimensions. For instance, to obtain comments of good quality, special training and clear guidance should be provided for evaluators to let them know what makes good or poor comments.

While numeric ratings seem to be dominating in organizations' performance appraisal practices, the role of narrative comments should not be overlooked. When we understand the importance of qualitative data, a question should also be asked: "How can we have people provide good comments in the evaluation?" This study intended to delineate a picture of personality traits and production of comments. After all, communication of performance feedback messages should pivot on the end user. The ultimate purpose is to provide them clear and useful information that will help them understand, reflect and improve. Narrative comments have their unique value as a supplementary to numeric ratings. More extensive research effort should be made about this area in performance appraisal.

Theoretical Contribution and Practical Implications

Due to the prevalence of using quantitative data in performance appraisal, a number of studies have converged on the application of numeric ratings (e.g. Bernardin et al., 2000; Kane et al., 1995; Kuang, 2005; Spence & Keeping, 2010; Yun et al., 2005). In

those studies, personality traits such as Conscientiousness or Agreeableness have often been linked to the variance in evaluation behaviors such as rating elevation or leniency. This research highlighted the significance of qualitative data that had been overlooked and made an attempt to look into the connection between individual characteristics and the production of comments. As comment construction is a process that consists of memory recollection, information organization and text production, relevant evidence was obtained from a variety of literature covering areas of management, psychology as well as linguistics for a thorough analysis.

It is believed that understanding the relationship between individual characteristics and production of comments has practical implications for practitioners. As introduced at the beginning of this research, since each individual is unique, the difference of personality traits is likely to be mirrored in their writing behaviors. How much, how specifically, or in what tone to write, tends to vary due to individual characteristics. However, as introduced earlier, in a performance appraisal context, narrative comments may serve multiple purposes. Practitioners use the comments as references for administrative purposes such as regular review of performance, promotion, or pay adjustment; employees review and reflect their work performance in the past as well as obtain suggestions for future improvement by reading the comments. To achieve these ultimate purposes, close attention should be paid to how to minimize the influences of personality traits and ensure the objectivity and quality of comments. On one hand, subjective comments may undermine the accuracy of evaluation outcome and will not help recipients to improve. For example, comments that are written in a positive tone do not necessarily represent real good performance of the evaluatee, but may just be a

consequence of the writer's agreeable personality. On the other hand, short and unspecific comments will not be informative and instructive to employees but become a redundant part of the evaluation. To use the data in this study as another example, comments like "I enjoy working with her" or "He is a smart guy" appear to be very tenuous as performance feedback and may fail to satisfy the needs of both practitioner and recipient.

Therefore, again, it is necessary to underline the importance of providing training and guidance prior to performance appraisals. Administrators should try to provide special trainings and clear guidance to the employees about how to write comments. Since writing comments requires cognitive effort in terms of recollecting information and text composition, the goal of training or guidance should be aimed at how to maximize such effort. It is advised that the evaluator spend a certain length of time in thinking about relevant performance cues about the evaluatee and writing the comment text. Also, instead of describing personality traits or personal feelings, the content of comments should focus on work-related behaviors or tasks. In particular, it was found in this study that males wrote comments focusing more on behaviors than females did. It suggests that it may be necessary to direct such trainings to female audiences helping them overcome their tendency to focus on personal feelings. Another suggestion would be to request the comment writers to provide specific numbers in comments to support their evaluation. Since one of the potential shortness of qualitative data is its subjectivity, it is believed that the presence of specific description of behaviors and concrete numbers can mitigate the potential influences of subjective factors and make the comments more convincing,

which enhances the feedback recipient's fairness perception of the evaluation and motivation for future improvement.

Also, it is necessary to realize that, due to the influences of individual characteristics on production of comments, qualitative data should be collected from a broader base or source of evaluators in the organization. A very good answer to this need is multi-source assessment, i.e. 360-degree feedback, in which evaluation data input is sought from an employee's supervisors, peers, subordinates, and even clients. From evaluation data that is sought from multiple sources, recipients will be able to receive more accurate information to reflect upon, and appraisal administrators will also be able to extract evaluation outcome more thoroughly.

References

Ang, S., Dyne, L. V., & Koh, C. (2006). Personality corrrelates of the four-factor model of cultural intelligence. *Group & Organization Management, 31(1),* 100-123.

Antonioni, D. (1996). Designing an effective 360-degree feedback appraisal process. *Organizational Dynamics, 25,* 24-38.

Baker, T. J., & Bechsel, J. (2006) Personality predictors of intelligence: Differences between young and cognitively healthy older adults. *Personality and Individual Differences, 41,* 861-871.

Barrick, M. R., & Mount, M. K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology, 44*, 1-26.

Berge, Z. (1997). Computer conferencing and the on-line classroom. *International Journal of Educational Telecommunications, 3(1),* 3–21.

Bernardin, H. J., Cooke, D. K., & Villanova. P. (2000). Conscientiousness and agreeableness as predictors of rating leniency. *Journal of Applied Psychology, 85,* 232-236.

Bidjerano, T., & Dai, D. Y. (2007). The relationship between the big-five model of personality and self-regulated learning strategies. *Learning and Individual Differences, 17,* 69–81.

Bradner, E., Mark, G., & Hertel, T. D. (2005). Team size and technology fit: Participation, awareness, and rapport in distributed teams. *IEEE transactions on Professional Communication, 48(1),* 68-77.

Bretz, R., Milkovich, G. T., & Read, W. (1992). The current state of performance appraisal research and practice: Concerns, directions, and implications. *Journal of Management, 18,* 321-352.

Brutus, S. (2010). Words and numbers: A theoretical exploration of giving and receiving narrative comments in performance appraisal. *Human Resource Management Review, 20(2),* 144-157.

Caldwell, D. F., & Burger J. M. (1998). Personality characteristics of job applicants and success in screening interviews. *Personnel Psychology, 51(1),* 119-136.

Callahan, J. L., Hasler, M. G., & Tolson, H. (2005). Perceptions of emotion expressiveness: Gender differences among senior executives. *Leadership & Organization Development Journal, 26(7/8),* 512-528.

Carment, D. W., Miles, C. G., & Cervin, V. B. (1965). Persuasiveness and persuasibility as related to intelligence and extraversion. *British Journal of Social and Clinical Psychology, 4,* 1–7.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112(1),* 155-159.

Costa, P. T., & McCrae, R. R. (1992). *Revised NEO Personality Inventory and the NEO Five-Factor Inventory.* Odessa, FL: Psychological Assessment Resources.

Daft, R. L., & Lengel, R. H. (1984). Information richness: A new approach to managerial behavior and organizational design. In L. L. Cummings & B. M. Staw (Eds.), *Research in organizational behavior.* (pp. 191-233). Homewood, IL: JAI Press.

Daft, R. L., & Lengel, R. H. (1986). Organizational information requirements, media richness, and structural design. *Management Science, 32(5),* 554-571.

Delassio, A. T. (1998). Using multisource 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.

De Raad, B., & Schouwenburg, H. C. (1996). Personality in learning and education: A review. *European Journal of Personality, 10,* 303–336.

Digman, J. M. (1990) Personality structure: Emergence of the five-factor model. *Annual Review of Psychonology, 41,* 417-440.

Digman, J. M., & Inouye, J. (1986). Further specification of the five robust factors of personality. *Journal of Personality and Social Psychology, 50(1),* 116–123.

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.

Fisher, C. D. (1989). Current and recurrent challenges in HRM. *Journal of Management, 15,* 157-180.

Fletcher, C. (1999). The implication of research on gender differences in self-assessment and 360 degree appraisal. *Human Resource Management Journal, 9(1),* 39-46.

Fong, E. A., & Tosi, H. L. Jr. (2007). Effort, performance, and conscientiousness: An agency theory perspective. *Journal of Management, 33(2),* 161-179.

Fried, Y., Tiegs, R. B., & Bellamy, A. R. (1992). Personal and interpersonal predictors of supervisors' avoidance of evaluating subordinates. *Journal of Applied Psychology, 77,* 462-468.

Furnham, A., & Stringfield, P. (2001). Gender differences in rating reports: Female managers are harsher raters, particularly of males. *Journal of Managerial Psychology, 16(4),* 281-288.

Geer, J. G. (1991). Do open-ended questions measure "salient" issues? *Public Opinion Quarterly, 55,* 360-370.

Gifford, R., & Hine, D.W. (1994). The role of verbal behavior in the encoding and decoding of interpersonal dispositions. *Journal of Research in Personality, 28,* 115–132.

Goldberg, L. R. (1990). An alternative "description of personality": The Big-Five factor structure. *Journal of Personality and Social Psychology, 59,* 1216-1229.

Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist, 48,* 26-34.

Grabe, W., & Kaplan, R. B. (1996). *Theory and practice of writing: An applied linguistic perspective.* New York: Longman.

Gray, J. R., & Braver, T. S. (2002). Personality predicts working-memory–related activation in the caudal anterior cingulate cortex. *Cognitive, Affective, & Behavioral Neuroscience, 2(1),* 64-75.

Gross, J. J., & John, O. P. (1997). Revealing feelings: Facets of emotional expressivity in self-reports, peer ratings, and behavior. *Journal of Personality and Social Psychology, 72(2),* 435-448.

Harman, R. P. (2009). Understanding survey comment nonresponse and the characteristics of nonresponders. *Dissertation Abstracts International, B70(01)*. (UMI No. 3345384)

Ilgen, D. R., Barnes-Farrell, J. L., & Mckellin, D. B. (1993). Performance appraisal process research in the 1980s: What has it contributed to appraisals in use? *Organizational Behavior and Human Decision Processes, 54,* 321-368.

Indartono, S., & Chen, C. H. V. (2010). Moderation of gender on the relationship between task characteristics and performance. *International Journal of Organizational Innovation, 2(2),* 195-222.

International Personality Item Pool, http://ipip.ori.org.

Jackson, K. M., & Trochim, W. M. K. (2002). Concept mapping as an alternative approach for the analysis of open-ended survey responses. *Organizational Research Methods, 5,* 307-336.

Jo, K. (1990). But what do they say? Gender and the content of student writing. *Discourse Processes, 13(2),* 243-257.

John, O. P. (1989). Towards a taxonomy of personality descriptors. In D. M. Buss & N. Cantor (Ed.), *Personality Psychology: Recent Trends and Emerging Directions.* (pp. 261-271). New York: Springer-Verlag.

Johns, G. (2006). The essential impact of context on organizational behavior. *Academy of Management Review, 31,* 386-408.

Jolson, M. A., & Corner, L. B. (1997). The use of instrumental and expressive personality traits as indicators of a salesperson's behavior. *Journal of Personal Selling & Sales Management, 17 (1),* 29–43.

Jones, S., & Kayongo, J. (2008). Identifying student and faculty needs through LibQUAL+™: An analysis of qualitative survey comments. *College and Research Libraries, 69,* 493-509.

Judge, T. A. & Ferris, G. R. (1993). Social context of job performance evaluation decisions. *Academy of Management Journal, 36,* 80-105.

Kane, J. S., Bernardin, H. J., Villanova, P., & Peyrefitte, J. (1995). Stability of rater leniency: Three studies. *Academy of Management Journal, 38,* 1039–1051.

Krugman, H. E. (1966). The measurement of advertising involvement. *Public Opinion Quarterly, 30 (4),* 583-596.

Kuang, D. C. Y. (2005). Assessing performance: Investigation of the influence of item context and rater personality. *Dissertation Abstracts International, 65(10-B).* (UMI No. 3150616)

Landy, F. J., & Farr, J. (1980). Performance rating. *Psychological Bulletin, 87,* 72-107.

Lavinia, M. B. (2004). A model for defining complexity in descriptive text type. *Folia Linguistica, 38(3-4),* 355-381.

LeBreton, J. M., & Senter, J. L. (2008). Answers to 20 questions about interrater reliability and interrater agreement. *Organizational Research Methods, 11(4),* 815-852.

LePine, J. A., Piccolo, R. F., Jackson, C. L., Mathieu, J. E., & Saul, J. R. (2008). A meta-analysis of teamwork processes: Tests of a multidimensional model and relationships with team effectiveness criteria. *Personnel Psychology, 61(2),* 273-307.

LePine, J. A., & Van Dyne, L. (2001). Voice and cooperative behavior as contrasting forms of contextual performance: Evidence of differential relationships with Big Five personality characteristics and cognitive ability. *Journal of Applied Psychology, 86,* 326-336.

Lieberman, M. D. (2000). Introversion and working memory: Central executive differences. *Personality and Individual Differences, 28,* 479-486.

Meyers-Levy, J., & Maheswaran, D. (1991). Exploring differences in males' and females' processing strategies. *Journal of Consumer Research, 18(1)*, 63-70.

Meyers-Levy, J., & Sternthal, B. (1991). Gender differences in the use of message cues and judgments. *Journal of Marketing Research, 28(1),* 84-96.

McCrae, R. R., & Costa, P. T. (1989). Reinterpreting the Myers-Briggs Type Indicator from the perspective of the five-factor model of personality. *Journal of Personality, 57,* 17-40.

Mehl, M. R., Gosling, S. D., & Pennebaker, J. W. (2006). Personality in its natural habitat: Manifestations and implicit folk theories of personality in daily life. *Journal of Personality and Social Psychology, 90(5),* 862-877.

Miller, N. M., O'Farrell, K., & DeMay, C. C. (2006, May). *Do comment characteristics help or hinder employee performance?* Paper presented at the 21st Society for Industrial Organizational Psychology Conference, Dallas, TX.

Mossholder, K. W., Settoon, R. P., Harris, S. G., & Armenakis, A. A. (1995). Measuring emotion in open-ended survey responses: An application of textual data analysis. *Journal of Management, 21,* 335-355.

Neuman, G. A., Wagner, S. H., & Christiansen, N. D. (1999). The relationship between work-team personality composition and the job performance of teams. *Group & Organization Management, 24(1),* 28-45.

Noftle, E. E., & Robins, R. W. (2007). Personality predictors of academic outcomes: Big Five correlates of GPA and SAT scores. *Journal of Personality and Social Psychology, 93,* 116–130.

Oberlander, J., & Gill, A. J. (2006). Language with character: A stratified corpus comparison of individual differences in e-mail communication. *Discourse Processes, 42(3),* 239-270.

Pearman, A. (2009). Basic cognition in adulthood: Combined effects of sex and personality. *Personality and Individual Differences, 47,* 357-362.

Pennebaker, J. W., Mehl, M. R., Niederhoffer, K. G. (2003). Psychological aspects of natural language use: Our words, our selves. *Annual Review of Psychology, 54,* 547–577.

Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin, 135(2),* 322-338.

Relja, V. (2007). On measuring language complexity as relative to the conveyed linguistic information. *SKY Journal of Linguistics, 20,* 399-427.

Rosen, L. D., Chang, J., Erwin, L., Carrier, L. M., & Cheever, N. A. (2010). The relationship between "textisms" and formal and informal writing among young adults. *Communication Research, 37(3),* 420-440.

Russ, G. S., Daft, R. L., & Lengel R. H. (1990). Media selection and managerial characteristics in organizational communications. *Management Communication Quarterly, 4(2),* 151-175.

Scherer, R. F., Owen, C. L., & Brodzinski, J. D. (1991). RATER AND RATEE SEX EFFECTS ON PERFORMANCE EVALUATION IN A FIELD SETTING A multivariate analysis. *Management Communication Quarterly: McQ(1986-1998), 5(2),* 174-191.

Schultz, B., & Anderson, J. (1986). Where do we know and where can we go? A summary and a response. *The American Behavioral Scientist(1986-1994), 29(3),* 368-376.

Schweiger, I., & Sumners, G. E. (1994). Optimizing the value of performance appraisals. *Managerial Auditing Journal: 9(8),* 3-7.

Selfe, C. L., & Meyer, P. R. (1991). Testing claims for on-line conferences. *Written Communication, 8(2),* 163-192.

Shore, L. M., & Thornton, G. C. III (1986). Effects of gender on self- and supervisory ratings. *The Academy of Management Journal, 29(1),* 115-129.

Sims, R. R., Veres III, J. G., & Heninger, S. M. (1987). Training appraisers: An orientation program for improving supervisory performance ratings. *Public Personnel Management, 16(1),* 37-46.

Smither, J. W., & Walker, A. G. (2004). Are the characteristics of narrative comments related to improvement in multirater feedback over time? *Journal of Applied Psychology, 89,* 575-581.

Spence, J. R., & Keeping, L. M. (2010). The impact of non-performance information on ratings of job performance: A policy-capturing approach. *Journal of Organizational Behavior, 31,* 587-608.

Stefanone, M. A., & Jang, C. Y. (2007). Writing for friends and family: The interpersonal nature of blogs. *Journal of Computer-Mediated Communication, 13,* 123-140.

Stewart, C. M., Shields, S. F., Monolescu, D., & Taylor J. C. (1999). Gender and participation in synchronous CMC: An IRC case study. *Interpersonal computing and technology, 7(1-2),* 1-25.

Sulsky, L. M., & Keown, J. L. (1998). Performance appraisal in the changing world of work: Implications for the meaning and measurement of work performance. *Canadian Psychology, 39,* 52-59.

Tannen, D. (1990). *You just don't understand women and men in conversation.* New York: Ballantine Books.

Tannen, D. (1994). *Gender and discourse.* New York: Oxford University Press, Inc.

Taylor-Nelms, L. (2003). Gender tales in computer-mediated communication. *Dissertation Abstracts International, 64(5).* (UMI No. 3091309)

Thoms, P., Moore, K. S., & Scott, K. S. (1996). The relationship between self-efficacy for participating in self-managed work groups and the big five personality dimensions. *Journal of Organizational Behavior (1986-1998), 17,* 349-362.

Trautwein, U., Ludtke, O., Roberts, B. W., Schnyder, I., & Niggli, A. (2009). Different forces, same consequence: Conscientiousness and competence beliefs are

independent predictors of academic effort and achievement. *Journal of Personality and Social Psychology, 97(6),* 1115-1128.

Tziner, A., & Kopelman, R. E. (2002). Is there a Preferred Performance Rating Format? A Non-psychometric Perspective. *Applied Psychology: An International Review, 51(3),* 479-503.

Tziner, A., Murphy, K. R., & Cleveland, J. N. (2002). Does conscientiousness moderate the relationship between attitudes and beliefs regarding performance appraisal and rating behavior? *International Journal of Selection and Assessment*, *10,* 218–224.

Tziner, A., Murphy, K. R., & Cleveland, J. N. (2005). Contextual and rater factors affecting rating behavior. *Group & Organization Management, 30(1),* 89-98.

Vignovic, J. A., & Thompson, L. F. (2010). Computer-mediated cross-cultural collaboration: Attributing communication errors to the person versus the situation. *Journal of Applied Psychology, 95(2),* 265-276.

Waldman, D. A., Atwater, L. E., & Antonioni, D. (1998). Has 360 degree feedback gone amok? *Academy of Management Executive, 12(2),* 86-94.

Whetten, D. A., & Cameron, K. S. (2005). *Developing management skills.* Upper Saddle River, NJ: Prentice Hall.

Yun, G. J., Donahue, L. M., Dudley, N. M., & Mcfarland, L. A. (2005). Rater personality, rating format, and social context: Implications for performance appraisal ratings. *International Journal of Selection and Assessment, 13(2),* 97-107.

Appendix A — Tables

Table 1

Intercorrelations Among Key Variables

Variable	N	*M*	*SD*	1	2	3	4	5	6	7	8
1. Amount of Comments	69	2.08	1.06	.98							
2. Specificity of Comments	69	2.72	1.16	.43**	.97						
3. Positivity of Comments	69	4.28	0.95	.09	.16	.97					
4. Conscientiousness	69	3.60	0.53	.10	.10	.08	.77				
5. Extraversion	69	3.56	0.62	.04	.12	.14	.34**	.87			
6. Agreeableness	69	3.77	0.45	.19	.14	.09	.43**	.26*	.71		
7. Group Size	69	3.71	0.82	-.13	.01	-.02	-.03	.15	-.04	–	
8. Gender (1=male, 2=female)	69	1.61	0.49	.07	.24*	-.13	.11	-.05	.16	-.03	–

Note: Reliability estimates are presented on the diagonal.

* $p < .05$. ** $p < .01$.

Table 2

Hierarchical Regression Analyses with Amount of Comments as Dependent Variable, Conscientiousness as Predictor, and Group Size as Moderator

	N	Model 1		Model 2		Model 3	
Variable		*B*	*SE B*	*B*	*SE B*	*B*	*SE B*
1. Gender	87	.27	.28	.26	.27	.25	.28
2. Conscientiousness (C)	87			.12	.27	.12	.27
Group Size (GS)	87			-.18	.16	-.19	.17
3. Two-way interactions (C * GS)	87					.09	.34
R^2		.01		.03		.03	
ΔR^2		.01		.02		.00	

Table 3

Hierarchical Regression Analyses with Specificity of Comments as Dependent Variable, Conscientiousness as Predictor, and Group Size as Moderator

	N	Model 1		Model 2		Model 3	
Variable		*B*	*SE B*	*B*	*SE B*	*B*	*SE B*
1. Gender	69	.58*	.28	.56	.29	.56	.29
2. Conscientiousness (C)	69			.16	.26	.16	.27
Group Size (GS)	69			.03	.17	.03	.17
3. Two-way interactions (C * GS)	69					.03	.36
R^2		.06*		.07		.07	
ΔR^2		.06*		.01		.00	

* $p < .05$.

Table 4

Hierarchical Regression Analyses with Amount of Comments as Dependent Variable, Extraversion as Predictor, and Group Size as Moderator

	N	Model 1		Model 2		Model 3	
Variable		*B*	*SE B*	*B*	*SE B*	*B*	*SE B*
1. Gender	87	.27	.28	.30	.28	.33	.28
2. Extraversion (E)	87			.20	.23	.19	.23
Group Size (GS)	87			-.20	.16	-.16	.17
3. Two-way interactions (E * GS)	87					.23	.24
R^2		.01		.04		.05	
ΔR^2		.01		.02		.01	

Table 5

Hierarchical Regression Analyses with Specificity of Comments as Dependent Variable, Extraversion as Predictor, and Group Size as Moderator

	N	Model 1		Model 2		Model 3	
Variable		*B*	*SE B*	*B*	*SE B*	*B*	*SE B*
1. Gender	69	.58*	.28	.59*	.28	.62*	.29
2. Extraversion (E)	69			.24	.23	.24	.23
Group Size (GS)	69			.00	.17	.05	.18
3. Two-way interactions (E * GS)	69					.19	.24
R^2		.06*		.08		.09	
ΔR^2		.06*		.02		.01	

* $p < .05$.

Table 6

Hierarchical Regression Analyses with Positivity of Comments as Dependent Variable and Agreeableness as Predictor

	N	Model 1		Model 2	
Variable		*B*	*SE B*	*B*	*SE B*
1. Gender	69	-.26	.23	-.29	.24
2. Agreeableness	69			.24	.26
Group Size (GS)	69			-.02	.14
R^2		.02		.03	
ΔR^2		.02		.01	

Appendix B — Participant Recruitment Letter

Dear fellow students,

I am a student from MSCA program of JMSB and I am seeking your voluntary participation in a research project. I am conducting a study concerning individual characteristics and performance appraisal results. More specifically, I am interested in how some personality characteristics are related to how you will evaluate your peers later this semester. The survey will take only 10-15 minutes to complete.

Your identification, i.e. your student ID, will be needed in order to link your personality traits to the results of the peer evaluation system that you will use at the end of class. Note that the information that you provide will NOT be disclosed to any third party and the research results will ONLY be used for research purposes. Your participation is critical to the success of my thesis and I sincerely hope you will help.

Your time is greatly appreciated.

Fang (June) Wang

Appendix C — Consent Form

CONSENT TO PARTICIPATE IN RESEARCH PROJECT

Individual Characteristics in Performance Appraisal

This is to state that I agree to participate in a research project being conducted by Dr. Stéphane Brutus and Fang Wang of MSCA program of Concordia University (Contact Information: fan_wan@jmsb.concordia.ca).

A. PURPOSE

I have been informed that the purpose of the research is to examine the relationship between individual characteristics and the results of peer evaluation that will be conducted at the end of the semester. More specifically, I will complete a short survey whose result will be linked with some of the information collected via the peer evaluation system.

B. PROCEDURES

The study will be conducted at John Molson School of Business at Concordia University. I will complete a brief survey and provide my student ID. This information is only needed to match my survey response with my response on the peer evaluation process. The information that I provide will NOT be disclosed to any third party and the research results will ONLY be used for the research purposes, in aggregate format.

C. RISKS AND BENEFITS

There is no risk of participation. I will be offered partial credits by participating in the study.

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 (i.e., the researcher will know, but will not disclose my identity)

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

NAME (please print)

_____

SIGNATURE

_____

If at any time you have questions about your rights as a research participant, please contact Dr. Brigitte Des Rosiers, Research Ethics and Compliance Officer, Concordia University, at (514) 848-2424 x. 7481 or by e-mail at bdesrosi@alcor.concordia.ca.

Appendix D — Questionnaire

We thank you for your cooperation in this project. If you are interested in learning about the results of this study, please do not hesitate to contact me (fan_wan@jmsb.concordia.ca) or my thesis supervisor Dr. Stéphane Brutus (brutus@ jmsb.concordia.ca). Once the study is complete, we would be pleased to send you a summary of the findings.

Please first provide the following information.

Student ID:

Gender:

Major:

Number of your project team members:

This survey is to test your personality traits. For each item below, please choose the number that is closest to your answer.

1— Very Inaccurate

2— Moderately Inaccurate

3— Neither Inaccurate nor Accurate

4— Moderately Accurate

5— Very Accurate

1. I am always prepared.

   1 2 3 4 5

2. I pay attention to details.

   1 2 3 4 5

3. I get chores done right away.

   1 2 3 4 5

4. I carry out my plans.

   1 2 3 4 5

5. I make plans and stick to them.

   1 2 3 4 5

6. I waste my time.

   1 2 3 4 5

7. I find it difficult to get down to work.

   1 2 3 4 5

8. I do just enough work to get by.

   1 2 3 4 5

9. I don't see things through.

   1 2 3 4 5

10. I shirk my duties.

1 2 3 4 5

11. I feel comfortable around people.

1 2 3 4 5

12. I make friends easily.

1 2 3 4 5

13. I am skilled in handling social situations.

1 2 3 4 5

14. I am the life of the party.

1 2 3 4 5

15. I know how to captivate people.

1 2 3 4 5

16. I have little to say.

1 2 3 4 5

17. I keep in the background.

1 2 3 4 5

18. I would describe my experiences as somewhat dull.

1 2 3 4 5

19. I don't like to draw attention to myself.

1 2 3 4 5

20. I don't talk a lot.

1 2 3 4 5

21. I have a good word for everyone.

1 2 3 4 5

22. I believe that others have good intentions.

1 2 3 4 5

23. I respect others.

1 2 3 4 5

24. I accept people as they are.

1 2 3 4 5

25. I make people feel at ease.

1 2 3 4 5

26. I have a sharp tongue.

1 2 3 4 5

27. I cut others to pieces.

1 2 3 4 5

28. I suspect hidden motives in others.

1 2 3 4 5

29. I get back at others.

1 2 3 4 5

30. I insult people.

1 2 3 4 5

Appendix E — Coding Protocol

Individual Characteristics in Performance Appraisal

Data Coding Protocol

Prepared by Fang (June) Wang
Email: fan_wan@jmsb.concordia.ca

General Introduction

The written comments to be coded are from peer evaluations given by COMM 222 students for their group projects. In this study, there are three variables to be analyzed about the comments: 1) amount of comments, 2) specificity of comments and 3) positivity of comments. We will both read the comments and manually code the variables independently. If you prefer working on a hard copy of the data, please use a pencil to make notes, which will be easier for any modification. The coding instructions will be as follows. If you have any questions about the instructions, please contact me before you start coding.

Instructions

1. Amount of comments

The amount of comments will be measured by the average number of "narrative units" in the comments written by a student. To be more specific, you can code step by step as follows:

Step 1: Mark and summarize each narrative unit, and then rate the number of narrative units in each comment (i.e. the comment for one team member). Here, "narrative units" specifically refer to those meaningful and distinguishable performance-related themes found in comments. A narrative unit can be multiple sentences, one sentence, one phrase or even one single word, as long as its content contains one and only one performance-related theme. Performance-related themes may be about behaviors or personal traits of a person.

Example 1: " May is not a good teammate. Every time we have a group meeting, she always arrives late and sometimes even doesn't show up. She doesn't talk much to any of us.[1]"

1- cooperation

In this comment, as underlined above, there is only one performance-related theme – which can be summarized as "cooperation" or "teamwork spirit". The first sentence

is a general evaluation, and the rest of the comment just explicates how May is not cooperative. This comment will be rated as *1* narrative unit.

Example 2: "He always contributes new ideas in the project[1] and delivers his part of work in time[2]."

1- idea contribution

2- practical contribution

In this comment, as underlined above, there are two performance-related themes – which can be summarized as "idea contribution" and "practical contribution", so it will be rated as *2* narrative units.

Example 3: "Overall she is a good team member.[1] I enjoy working with her.[2]"

1- teamwork spirit

2- personal feeling

This comment is rather vague and contains no specific performance-related theme. It provides a general evaluation about the teammate's performance and describes the rater's feeling of working with her. It will be rated as *2* narrative units.

Step 2: After rating the number of narrative units in each comment, add up the sum of narrative units in all the comments written by that student. Write down the result.

2. Specificity of comments

The specificity of comments will be measured by the degree of **trait versus behavior focus** in the comments. Here "trait" can be understood as a quality, ability or characteristics of a person. Comments that focus more on performance-related behaviors are more specific. You can code step by step as follows:

Step 1: Rate the specificity of each comment (i.e. the comment for one team member). A 5-point likert scale will be used, in which:

1 = comment that focuses entirely on personal traits or characteristics;

2 = comment that focuses mostly on personal traits or characteristics, but to a lesser extent also refers to specific behaviors or tasks;

3 = comment that focuses equally on personal traits or characteristics and specific behaviors or tasks, or one that focuses neither on traits nor behaviors;

4 = comment that focuses mostly on specific behaviors or tasks but to a lesser extent also refers to personal traits or characteristics;

5 = comment that focuses entirely on specific behaviors or tasks.

Example 1: "Jack has a great humorous sense and always makes us laugh. I like working with him."

This comment focuses entirely on personal trait. It will be rated *1*.

Example 2: “Jack appears a little bit shy and silent in the group meetings, but I can tell he’s actually the most hard-working one in the team. He writes a lot in the project report.”

This comment describes Jack a shy, silent but hard-working person, which are all his personal traits, and also mentions a specific behavior——writing. It will be rated *2*.

Example 3: “May has been very cooperative in the project. She always responds to our questions for her work quickly. However I think she needs to be more honest with us. At times she seems to give different answers to the same question.”

This comment describes two personal traits of May’s——cooperative and dishonest, each followed by an explanation that focuses on a specific behavior. It is considered as focusing equally on personal traits or characteristics and specific behaviors, so it will be rated *3*.

*You may find some comments which are so general that focus on neither personal traits nor specific behaviors. This kind of comments will also be rated 3.

Example 4: “He is an excellent team leader, by keeping us informed of progress and following up our feedback timely. In addition, he can come up with quick solutions to various problems.”

In this comment, “an excellent team leader” refers to a personal characteristic, and the rest focuses on specific behaviors. There are more descriptions of specific behaviors than that of personal characteristic, therefore, it will be rated *4*.

Step 2: After rating the specificity of each comment, add up the ratings of all the comments written by that student. Write down the result.

3. Positivity of comments

Step 1: You will read the comments and rate the positivity according to your judgment, using a 5-point scale in which 1 = very negative, 2 = slightly negative, 3 = neutral, 4 = slightly positive, 5 = very positive. In another word, within the 5-point scale, the higher the rating, the higher the positivity is. The rating should be given based on the degree of positive tone, no matter the comment is related to performance (behavior/task), personal trait, or any other dimensions.

Example 1: ““He always contributes new ideas in the project and delivers his part of work in time.”

Obviously, this comment only contains positive evaluation. It will be rated *5 - very positive*.

Example 2: “Jack has a great humorous sense and always makes us laugh. I like working with him.”

Similarly, this comment only contains positive evaluation, even though it is about Jack’s personality. It will be rated *5 - very positive.*

Example 3: “May likes to share her ideas in the meetings. Although some of them are not very realistic, I appreciate her active participating attitude. Personally I’d suggest her do more text readings before giving an idea”.

In this comment, it praises May’s participating attitude by actively sharing ideas (positive), and also points out that her ideas are sometimes unrealistic (negative), and then makes a suggestion in positive tone (positive). It will be rated *4 – slightly positive.*

Example 4: “May is not a good teammate. Every time we have a group meeting, she always arrives late and sometimes even doesn’t show up. She doesn’t talk much to any of us. Fortunately, at the end, she did finish her part in our final report in time and write well.”

This comment strongly criticizes May’s poor cooperation manner in teamwork, and only gives one positive evaluation about her writing in the report. Overall, the comment is in negative tone. It will be rated *2 – slightly negative.*

Step 2: After rating the positivity of each comment, add up the ratings of all the comments written by that student. Write down the sum.