Clemson University
TigerPrints

All Theses Theses

8-2007

Standardized Power: A Rhetoric of Performance Evaluation in Education

Michael Hedges
Clemson University, mnhedges@yahoo.com

Follow this and additional works at: https://tigerprints.clemson.edu/all_theses

Part of the Rhetoric and Composition Commons

Recommended Citation
Hedges, Michael, "Standardized Power: A Rhetoric of Performance Evaluation in Education" (2007). *All Theses*. 166.
https://tigerprints.clemson.edu/all_theses/166

This Thesis is brought to you for free and open access by the Theses at TigerPrints. It has been accepted for inclusion in All Theses by an authorized administrator of TigerPrints. For more information, please contact kokeefe@clemson.edu.

STANDARDIZED POWER: A RHETORIC OF PERFORMANCE EVALUATION IN EDUCATION

A Thesis
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Master of Arts
Professional Communication

---

by
Michael N. Hedges
August 2007

Accepted by:
Dr. Cynthia Haynes, Chair
Dr. Tharon Howard
Dr. Jan Holmevik

ABSTRACT

South Carolina's public education system administrators and teachers need to know more about how the language used in written performance evaluations impacts the effectiveness of evaluation feedback in order to help improve the state's system for Assisting, Developing, and Evaluating Professional Teaching (ADEPT). One key to understanding this knowledge is identifying the characteristics of the language used in evaluation feedback. This study examines the rhetorical situation of written performance evaluations from a theoretical standpoint and used a survey to assess several characteristics of the language used in evaluation feedback (word usage, content inclusion, and the structural order of sentences). Results, while tentative, indicate that effective written performance evaluation feedback avoids negative framing, demonstrates rater awareness of the performance situation, and avoids feedback in question form. The implications of the rhetorical theory behind evaluation and the study's findings are discussed for the fields of professional communication, industrial/organizational psychology, and education. Future directions for research and practice are proposed.

DEDICATION

This work is dedicated to my wife, Jamie, whose passion for teaching ignited my own. It is also dedicated to all those who teach, formally or informally, and those who have inspired and encouraged their endeavors along the way.

ACKNOWLEDGEMENTS

I thank everyone who has been a part of this long process. Most of all, I am enormously appreciative of the direction I had under Dr. Cynthia Haynes. Dr. Haynes' guidance, patience, persistence, and enthusiasm made this work enjoyable despite setbacks of all varieties and kept my confidence high. I am truly appreciative of Dr. Tharon Howard, who guided me through the Master of Arts in Professional Communication and served as a reader on this thesis. His challenging dialogue made this work more complete. I am also truly appreciative of Dr. Jan Holmevik, who provided valuable insight as a newcomer to our academic field and also rescued a desperate student from quite a pickle. Finally, I am grateful for the initial contributions of Dr. Steve Katz. It is also prudent to mention Dr. Patrick Raymark, who sparked my interest in performance evaluation. Thank you to all the faculty, friends, and fellow students who have served as ears to listen as I worked out my thoughts regarding this paper.

TABLE OF CONTENTS

Page

TITLE PAGE i

ABSTRACT iii

DEDICATION v

ACKNOWLEDGEMENTS vii

LIST OF TABLES xi

LIST OF FIGURES xiii

CHAPTER

1. INTRODUCTION 1

2. METHODOLOGY 31

3. RESULTS 43

4. CONCLUSIONS AND FUTURE DIRECTIONS 61

APPENDICES

A. Pilot Study Findings 79

B. Pilot Study Questionnaire 87

C. Recruitment Letter 91

D. Survey Invitation 93

E. ADEPT Information 95

F. Online Survey 109

REFERENCES 115

LIST OF TABLES

Tables Page

1. Results of Survey Item: "You failed to manage the classroom properly because you did not maintain authority." 44

2. Results of Survey Item: "You might improve your classroom management by maintaining authority." 45

3. Results of Survey Item: "Because you did not maintain authority I have never seen poorer classroom management." 46

4. Results of Survey Item: "One way to improve classroom management would be to maintain authority." 47

5. Results of Survey Item: "Try other methods to achieve classroom management." 49

6. Results of Survey Item: "While doing your lesson plan what were your accommodations." 50

7. Results of Survey Item: "Good way to quiet down your kids!" 51

8. Results of Survey Item: "Nice way to offer a review of material." 52

9. Results of Survey Item: "Why did you discipline student A that way?" 54

10. Results of Survey Item: "You handled your class well. Perhaps next time you will consider disciplining student A by…." 55

11. Results of Survey Item: "You did not discipline student A appropriately. But you handled the class well." 56

12. Results by Gender of Survey Item: "You failed to manage the classroom properly because you did not maintain authority." 57

13. Results by Gender of Survey Item: "You might improve your classroom management by maintaining authority." 58

14. Results by Gender of Survey Item: "You handled your class well. Perhaps next time you will consider disciplining student A by…." 59

LIST OF FIGURES

Figure Page

1. Words Considered Counterproductive to Improvement 48

2. Open Response to Performance Scenarios 53

3. Actual Perception of Contextual Awareness Between Ratee and Rater 65

4. Ideal Perception of Contextual Awareness Between Ratee and Rater 66

CHAPTER ONE

INTRODUCTION

Mr. Lee is a second year English teacher at a public high school. Today he will be evaluated according to the state mandated performance evaluation system that is designed to improve teaching in the classroom. He is well prepared to teach. He always prepares lesson plans well in advance, tailors assignments to students' needs, and uses his knowledge base as a candidate for a Master's degree to make sure that the curriculum is informative and interesting. Currently, he teaches two sections of basic sophomore level English and one pre-Advanced Placement (A.P.) English course. Mr. Lee is aware of the performance evaluation system. He has been evaluated before, but has always been unclear of the system's purpose and logic. However, he does know that poor evaluations reduce his chances of obtaining a continuing contract to teach in the school district.

Administrator Johnson is on the way to observe Mr. Lee's classroom. Johnson is a 20 year veteran of the public education system, having taught middle school mathematics for four years prior to becoming an assistant principal at the high school. Administrator Johnson's other duties include attendance enforcement and student parking management, but all administrators at the high school also must conduct evaluations of teachers as a requirement of the state's evaluation system. Administrator Johnson was instructed how to conduct evaluations under the current system two years ago during a three day seminar. Mr. Lee is Johnson's third evaluation of the day.

Mr. Lee begins class as Administrator Johnson observes him. The class is in the process of reviewing previous chapters of an assigned book when T.J. enters the classroom. T.J. has been placed in sophomore English for the third time, largely because of repeated trips to alternative school and suspensions that have resulted in incomplete coursework. T.J. received a referral, which is a disciplinary note sent to the principal, in Mr. Lee's class yesterday for throwing a pencil across the room. T. J., still upset over the referral, disrupts the class during the review and makes a remark directed at Mr. Lee that the class finds humorous. Mr. Lee loses ten minutes of instructional time while sending T.J. to in-school suspension and regaining students' attention. He is flustered; this is his third incident involving T.J. and the situation doesn't seem to be improving. Finally, the class continues on as Mr. Lee intended.

After the class concludes, Mr. Lee is concerned how the disruption might affect his evaluation. Administrator Johnson is not familiar with T. J. or his problematic background and has only observed Mr. Lee one other time during his pre-A.P. course. Administrator Johnson is frustrated. He must conduct the evaluation according to the framework of the assessment form, which simply requires to check off whether or not teachers are performing adequately on dimensions like meeting state curriculum standards, adapting lesson plans to student needs, and classroom management. The dimension descriptions leave Administrator Johnson no option but to give Mr. Lee poor marks regarding classroom management. Pressed for time and without a prompt for giving Mr. Lee encouragement and suggestions for how to handle problem students, Administrator Johnson finishes the evaluation and moves on to other duties. Mr. Lee is

left frustrated. He has no opportunity to explain the uniqueness of the incident and does not know how to improve. However, he does fear that receiving poor marks may affect his ability to continue his employment and becomes resentful toward the district and Administrator Johnson because he is powerless in the situation. Similarly, Administrator Johnson realizes that the evaluation system does not work and that he, too, holds little power over the process because of the system's constraints. Trapped within the system, both individuals begin to devalue evaluations while becoming more and more averse to the process. However, this situation is not unique to these two fictionalized individuals. Educators across America's public education systems are grappling how to make sense of evaluating teaching performance in an era of standardized practices.

The Impetus. The performance evaluation has long been the object of study by Industrial/Organizational (I/O) psychologists and the bane of the workplace for most employees, including teachers. From self-help books to the comic strip *Dilbert*, it is evident that the performance evaluation plays an integral part in many individuals' lives. However, there has not been enough exploration of the effects of performance evaluations on individuals being evaluated in contexts, such as teaching, where performance is not completely itemized or standardized. I/O psychology has played a major role in the present understanding of performance appraisal. For instance, Murphy and Cleveland (1995) conduct a comprehensive review of performance appraisal systems and examine standards for performance, obtaining performance information, error and accuracy measures and so forth. However, the review also demonstrates that I/O psychologists have not had enough scholarly discussion regarding the written feedback

that employees receive after evaluation. This is, perhaps, because not all performance evaluations rely upon written feedback from the evaluator in order to assess performance. But there are several occupations where the nature of the work dictates that specific commentary and advice on performance is necessary in order for the performer to improve. In order to better understand the written feedback that is sometimes a part of performance evaluation we might approach evaluations from a different theoretical perspective that examines the texts of performance evaluations and their surrounding rhetorical situation.

My own interest in performance evaluation practices arose from two exigencies. First, an I/O psychology course on the subject sparked interest and the material, or lack thereof, suggested that investigating evaluations from the perspective of a professional communicator might lead to a better theoretical and practical understanding of the mechanics and outcomes of the evaluation process. Discussions with I/O faculty and students and a review of current literature suggested strongly that the field had not given due attention to the importance of language. The field had inadvertently neglected to assess and acknowledge language's ability to radically impact the performance of the worker being evaluated and thus the organization's effectiveness. The text of performance evaluations was, in other words, largely assumed transparent and innocuous. The second exigency was a personal experience involving a secondary school teacher working in South Carolina's public education system that demonstrated to me that feedback in evaluations where language is not carefully considered does impact an employee's performance ability in ways that hinder performance. The state's system for

teacher evaluation, "Assisting, Developing, and Evaluating Professional Teaching" (ADEPT), was not conducive to professional development. The problem with the lack of attention to language is that we may have missed a crucial aspect of how to make performance evaluations more beneficial to the employee and employer and, consequently, improve workplace performance and the organizational climate. Clearly, this neglect of the rhetoric of written performance evaluations and their inherent rhetorical situation position professional communication as a field that can contribute to this issue.

Understanding Performance Evaluation

Formal performance evaluations are now an everyday practice in the workplace, but this was not always so. As Archer North and Associates, one of the world's more prominent performance appraisal system design companies explain, modern performance evaluations did not begin to occur until after World War II (*Introduction to Performance Appraisal*). At this time, standardization of business practices, especially work that resulted in a finished good or product, became more important. Performance evaluations were thus considered a way to control quality and maintain consistency. However, as Dulewicz (1989) explains, there seems to have been a human impetus to make judgments about one's self and others. From the time that little children race to see who is fastest, to grown men in a pick-up basketball game, we often measure ourselves against others in a variety of contexts. It is natural that these judgments and desire to measure would venture into the areas we are most familiar with: our work. Thus, performance

evaluations were implemented and designed within organizations in order to maintain a consistency of judgment. The idea is that this would make evaluation accurate and fair.

Originally, the purposes of evaluation were to measure company productivity and to make administrative decisions regarding employees. An individual's score on a yearly evaluation would determine whether or not that individual should receive a raise, maintain her current level of pay, or be fired. Evaluations that serve those purposes are commonly known as *summative*, because they are cumulative assessments of work performance that do not direct the employee on how to improve performance. These summative evaluations have little use for feedback directed at the employee, since they are only used by supervisors and such. Because these evaluations are out of employees' hands as soon as they are completed, there is a sense of finality to summative evaluation.

However, my research is primarily directed toward another type of evaluation: *formative*. It is now understood that if employees are made aware of their performance level, they may monitor their performance in order to improve. The purpose of formative evaluations is to not only assess an individual's current performance level but also to provide suggestions for how to better perform one's work duties. Formative evaluations imply a sense of on-going growth and development, in that their intent is for the employee to become a better worker.

The Impact of Evaluation

The majority of concerns with formative evaluations arise from the various proximal and distal effects of ineffective and negative evaluation feedback. For teachers,

those types of feedback may cause immediate and long term problems that interfere with their ability to perform their duties. They may find themselves with a lower internal motivation to teach immediately after an evaluation. This is because ineffective feedback may indicate to the evaluatee that performance is inadequate and unchangeable and that they are poor employees. In the long term, this lack of motivation not only leads to reduced classroom enthusiasm, but also a reluctance to investigate and implement different teaching strategies because performance is seen as static. However, the hesitancy to implement new strategies is even more problematic if the evaluatee was actually using incorrect teaching practices. The end result is that an evaluation designed to assess and improve performance has instead sentenced performance to the status quo at best or, more likely, actually hindered future performance because of inappropriate feedback.

It is also clear that ineffective or negative feedback in evaluations can lower self-esteem (Murphy & Cleveland, 1995). Low self-esteem in the workplace may cause poor performance. Teachers may feel inadequate or incapable as a result of feedback. This may lead to 1) instructional errors, 2) inappropriate relationships with students, and 3) improper attitudes about the teacher's self, such as an inflated ego as a compensation tactic. Low self-esteem may also lead to ineffective classroom management because of a teacher's lack of confidence to discipline students. Additionally, there is a direct relationship between mood and self-esteem at work and at home (Fuller et. al., 2003). The negative effects of poor feedback may travel well beyond the classroom. It can harm

personal relationships and lower the desire of an individual to participate in home life and leisure activities that might offset the apparent failures at work.

Research has found that the effect of written texts, such as the feedback on evaluation forms, on people's psyche can be much more powerful than one might think. Part of the aversion toward performance evaluation feedback may be a result that it makes the reader feel as if their identity is under attack. As Holland (1975; 1979) explains, readers interpret texts within their own selves: their backgrounds, their key life events, and their emotions shape meaning. While Holland dealt primarily with literature, it is suitable to think that readers bring a similar identity set to other written texts, including performance evaluations. Thus, feedback is manipulated by the reader until it is acceptable to his identity and rejected if it is unacceptable. This means that even the most well-meaning advice to correct an error may appear as a serious attack against the self.

But ineffective feedback in evaluations has significant impact well beyond the immediate realm of the evaluatee. It can contribute to the existing power differential between instructors and the administrators that evaluate them which may lead to a less cooperative and supportive workplace. As Bleich (1995) explains, traditional evaluation systems espouse the ideology that, despite being intimately involved in their work, teachers do not know best how to improve their work. This ideology is counterproductive as it implies a false sense of authority, placing administrators above teachers within the teaching context. Even when administrators do not personally endorse this ideology, they are victims of the system's ability to sublimate individual

convictions. Thus, this power differential is often as much unwanted by both parties as it is unavoidable.

M. Rose (1996) explains teachers' reluctance toward evaluation from a different perspective. He argues that current evaluation practices reinforce a negative stereotype of America's public education systems. Rose, critiquing the current climate surrounding education claims that "[w]e are offered, by both entertainment and news media, depictions of schools as mediocre places, where students are vacuous and teachers are not so bright…" and that this "engenders a mood of cynicism and retrenchment, preparing the public mind for extreme responses: increased layers of testing and control…" (1-2). Rose goes on to conclude, after a lengthy ethnographic study that covered public education teachers from coast to coast, that standardized evaluation falls short because there is no single profile of excellence in teaching. He recommends "no final list of good practices, no curricular framework or set of instructional guidelines" because "[s]uch profiles have a tendency to be stripped of context, to become rigid prescriptions, at times reduced to slogan or commodity" (9). Teachers, then, may be weary of evaluation because it trivializes their work and reinforces a misinformed stereotype.

Additionally, a distaste for evaluations could be one contributing factor to turnover in the education community, which is on the rise. For instance, it is possible that in the current academic year the South Carolina public education system will have more mathematics teachers retire or leave than they will certify new mathematics teachers (Hedges, J., personal communication, Aug. 24, 2006). In a passionate profession that traditionally offers many more intrinsic rewards than financial incentive it

makes little sense to enter a workplace that uses methods that hinder employees' ability and desire to perform.

Ultimately, all students' formal education is dependent upon the teacher. When something happens that negatively affects a teacher's ability to perform, it directly reduces the potential for learning. The stipulations of the No Child Left Behind Act also create consequences for inadequacy in the classroom. As L. Rose (2004) explains, schools that fail to meet academic progress rates for two consecutive years face several penalties including: 1) loss of funding, 2) termination of staff, and 3) dissolution of the school district. The long term result of inadequate teaching (whatever the cause) in the classroom, then, is quite grave: students are less prepared to continue their education or enter the working force.

These exigencies led me to further explore performance evaluation feedback with a pilot study (see Appendix A) which examined how school teachers react to evaluation feedback. The pilot study's findings were the impetus for the present study.

The Present Study

Clearly, in practice, evaluation feedback must be studied and assessed in order to improve the evaluation process and avoid harming educational systems' teachers, administrators, and students. Knowing appropriate words and phrases to use in performance evaluations should create a positive work experience for both the individual evaluating and the individual being evaluated. This will make constructive criticism toward the teacher more acceptable to that teacher, increasing the likelihood of

developmental change that results in teaching improvement. It should also increase evaluators' satisfaction with the appraisal process, as they may have increased confidence that their guidance will be fruitful. The subsequent improvement in teaching and teacher/administrator relationships may then may lead to better classrooms and better educated students.

Thus, this study proposes to identify *what administrators AND teachers consider as effective communication (which constitutes both appropriate use of rhetoric and knowledge of the rhetorical situation) in written performance appraisals through a survey administered to public school faculty who are receiving evaluations and performing evaluations*. For the purposes of this study:

1) *Administrators* are those individuals employed by a school system that oversee teachers and conduct performance appraisals of teachers.

2) *Teachers* are those individuals employed by a school system and are responsible for educating students.

3) *Effective communication* refers to written responses on the evaluation form that have the potential to produce developmental change.

4) *Performance appraisals* constitute an administrator observing a teacher in the classroom in order to assess the teacher's work-related competencies. A written form is attached to the evaluation in order to provide feedback to the teacher.

5) *Public school* refers to all levels of the public education system, which usually houses students from kindergarten through twelfth grades.

6) *Rhetoric* is the art of finding the available means of persuasion through speech

or writing in any given case.

7) *Rhetorical situation* refers to the context of a rhetorical act, and has three primary components: a) a rhetor, b) an issue, and c) an audience. In the case of this study the rhetorical act is a piece of writing that is delivered to the reader (rate or evaluatee) from the writer (rater or evaluator). These three components (text, reader, writer) form the rhetorical triangle that is inherently examined.

8) *Contextual awareness* refers to the degree of understanding and prior knowledge of a given situation that an individual brings to an evaluation scenario. More specifically, this study discusses contextual awareness in terms of both the evaluator and evaluatee. For the evaluator, contextual awareness is prior knowledge of the classroom being observed: its students, instructional goals, the teacher, and school activities that may disrupt or otherwise impact the class. For the evaluatee, contextual awareness is knowledge of the responsibilities and constraints that the evaluator brings to the observation. Particularly, this constitutes a knowledge of the standardized evaluation process and its goals.

The intent of the present study is to provide answers concerning the characteristics of the appraisal genre, the characteristics that raters and ratees in a particular discourse community value, and whether raters and ratees have similar perceptions regarding the genre. In particular, it will test the following three hypotheses:

Hypothesis 1

Teachers are more likely to identify ineffective word usage in evaluations than administrators. There will be disparity between administrator and teacher in this component of rhetorical awareness. Administrators might not be as aware of damaging words as teachers because administrators typically conduct evaluations much more than they receive them. Ineffective word usage will thus largely be apparent to teachers only because administrators have little understanding of the manner and context in which teachers receive evaluation comments.

Hypothesis 2

There is a strong positive correlation between contextual understanding and effective language in performance evaluation remarks. However, administrators and teachers do not share the same level of contextual understanding, leading to discrepancies in the perceived effectiveness of feedback. It is expected that a number of external factors such as time, interest, and knowledge of the appraisal process prevent administrators and teachers from sharing contextual understanding to the extent that content in evaluations always appears informed. Because of this phenomenon, it is believed that administrators will not value qualified statements that demonstrate contextual understanding as highly as teachers because, in practice, they are often accustomed to making comments with little prior information regarding the ratee's background and the current situation in which he is being observed.

Hypothesis 3

Neither administrators (raters) nor teachers (ratees) will consider negative statements in question form as effective communication. Although the pilot study demonstrated that raters occasionally make negative statements in question form, it is assumed that the accusatory nature of these statements will be perceived as inappropriate by both parties.

Understanding how raters and ratees use and perceive language in written performance appraisals will achieve a variety of goals across disciplines. First, it should increase awareness of how a performance appraisal effectively or ineffectively meets its goal and maintains its purpose- a question long considered by psychologists. It will also identify ways in which raters' and ratees' understandings of the appraisal process differ and how this evidences itself in the end product of evaluation.

Additionally, it provides an opportunity to analyze an unexplored genre, and will help professional communicators uncover its key conventions and how the genre is used. This should enable writing consultants tasked with designing evaluation instruments to be of additional benefit to their clients (Bergland, 1997; Ornatowski, 1995). More specifically, it will increase the ability to determine the validity of teacher evaluations, an area in which the education discourse community has expressed concern (Marshall, 2005; Tucker et. al, 2003).

The remainder of this chapter examines the research and current trends of the three academic fields (I/O psychology, education, and professional communication) that inform this study and constitute its theoretical framework. Chapter Two explores the

methodology utilized in the study and its limitations and degree of validity. Chapter Three is a dissemination of the study's findings and examines each of the three identified characteristics (content, word usage, and structure) as well as other aspects of the data that are relevant to the topic. Finally, Chapter Four discusses the conclusions regarding language in performance evaluations that may be inferred from the established theoretical perspective and the study's results. Then implications for practice and future research are explored.

But first, in order to provide rhetorical context and a theoretical framework, it is prudent to discuss the current state of evaluations in education and performance evaluation as it relates to the education, I/O psychology, and professional communication discourse communities.

The Current State of Evaluations in Education

Evaluations in school systems have been a hotly debated topic for the past decade. There are a great deal of disparate theories and opinions regarding how performance evaluations should be conducted, when they should occur, and who should do the evaluating. Much of this debate is a result of the standardization of teaching performance and its corresponding assessment instruments. This trend in evaluation may be linked to the emphasis on standardized testing that has influenced the nation's public education systems since the 1980's. Federal and state legislation now, more frequently than not, dictate the curriculum and method of measurement in the classroom.

Complicating this situation is the fact that the results of standardized testing are tied to financial incentives. It is in a school's best interest to have their students do well on the required tests because these schools receive rewards with which to improve basic and supplementary programs (i.e. multimedia textbooks and parent-student breakfasts). Additionally, the financial incentives are often established as rewards for *continuing* improvement. For example, a school might score in the 95th percentile in one academic year after being in the 87th percentile the previous year. This is an admirable achievement; however, the degree of improvement that the school demonstrated dictates that the reward level for the next year will be higher. In other words, if the school scores in the top five percent of all schools again, it may not meet the level required for a financial reward despite having another outstanding year.

The net effect of the current evaluation and reward system is that schools are left with fewer positive consequences from continued success and improvement and much to lose from performing poorly (loss of financial incentives, disruption of school's long term plans). This translates into a standardized, high risk classroom where teachers are required to "teach the test" in order to insure their school's livelihood. Thus, because of the stakes of standardized testing, teaching performance is more readily associated with the product (high test scores) than it is with the process of developing thinkers and learners.

The problem of teaching as product has significant impact upon the method and manner in which teachers are evaluated. Effective teaching is often quantified and categorized into a set of standards much like students' knowledge of academic subjects.

This has led to the implementation of performance evaluations that follow a standardized model. There is little room for extenuating circumstances and creative practices. It is as if educators are evaluated much like one would assess the value of a soft drink machine: does it dispense a beverage after being paid, and is that beverage the same beverage each time a button is pushed and a transaction occurs? Imagine the furor of a soft drink company if it discovered that one of its machines was dispensing a rival beverage that the machine thought might be of greater benefit to its customers. As abstract as this analogy appears, it is not unlike the climate of the public education system classroom. Administrators who evaluate teachers are doing so within the context of legislated, standardized testing, and this is apparently leading to a standardization of "right" and "wrong" teaching practices, with little room for teaching's organic nature in which learning occurs through a multitude of factors that differ from student to student based on abilities and learning styles.

Clearly, standardized testing will not disappear anytime soon because of our reliance upon the formal testing of students in order to measure ability. The benefits and risks associated with achievement and failure of standards will continue to dictate compliance with legislated standards. But even if it is impossible to de-standardize what is taught and how it is evaluated, it is possible to de-standardize the practical, tangible feedback that might encourage a teacher on toward better performance and a more rewarding work experience. For instance, Kyriakides, Demetriou, and Charalambous (2006) lament the lack of a theoretical framework for evaluation and charge that teachers themselves should be responsible for generating criteria for effective teaching practices.

Understanding South Carolina's ADEPT system. The state's system for Assisting, Developing, and Evaluating Professional Teaching (ADEPT) might best be understood using a hypothetical example that illustrates the evaluation process from start to finish (for the state provided explanation of the ADEPT system, see Appendix E). Mr. Lee, the teacher in the opening of this chapter, must go through four stages before he has completed the ADEPT process. These stages correspond with the program's acronym, albeit out of order. First, Mr. Lee goes through the *developing* stage. He would be assigned a mentor (generally, an administrator assigned by the school district) tasked with coaching Mr. Lee toward meeting the 10 performance dimensions that constitute an evaluation. Mr. Lee's mentor determines by the end of the academic year whether or not he is prepared to move to the next stage of the process, known as *assisting.* In the second year of Mr. Lee's career he will continue to be evaluated, according to the state standards, by a different assigned evaluator and receive assistance as needed. This assistance is generally limited to one mock evaluation earlier in the year because of the high workload of administrators.

At the conclusion of Mr. Lee's second year the process's focus shifts from improving Mr. Lee's teaching to assessing his teaching competence. He will go through the *evaluating* stage (even though Mr. Lee has been "evaluated" several times by now) and at this juncture the evaluation becomes high stakes. Mr. Lee's evaluator is now tasked with assessing whether or not Mr. Lee should receive a continuing contract to teach, which is similar to a professor receiving tenure in academia. In order to achieve a continuing contract, Mr. Lee must meet at least nine of the 10 performance dimensions.

The fourth phase, *professional teaching*, consists of goal-oriented evaluation in order to promote continuing professional development. In this phase, Mr. Lee must task *himself*, after gaining administrator approval, with what he needs to do (training seminars, graduate courses, correspondence programs, and the like) in order to improve his teaching performance. Mr. Lee will remain in this phase throughout his teaching career, readdressing and setting new goals yearly.

However, the above scenario assumes that Mr. Lee meets all requirements to move from one stage to the next. If Mr. Lee were not deemed ready by his mentor to move to the second stage of the program, he receives another year of development. Once Mr. Lee is ready for the administrative evaluation that determines his future employment, he receives two chances to pass. After the first failure, Mr. Lee must complete another year of teaching, receiving administrative assistance and evaluation, before retaking the formal evaluation. If he does not pass the second time he loses his certification to teach in the state of South Carolina for a period of two years. A remediation course must be passed in order for certification to be reissued and Mr. Lee must begin again at stage one of the process. If Mr. Lee were to fail the final evaluation at the end of this journey through the process he would be permanently banned from teaching in the state of South Carolina. However, Mr. Lee would still be eligible to teach in many other states, which do not have as strict regulations for retaining employment.

The ADEPT system's intent is an admirable one. But once it is understood how the process works several issues are apparent. Many of these issues are rooted in the two major purposes of evaluation (summative and formative) and that ADEPT tries to

accomplish both purposes with one evaluation program. The dual purposes may appear quite confusing to teachers and those who evaluate them. The evaluation instrument and its performance dimensions are the same regardless of the intent of the evaluation (for a more thorough critique of the evaluation instrument, see Chapter Four). Additionally, all evaluations are referred to as "formal" even if they are purely mock or developmental in nature. This may cause evaluators to not provide appropriate feedback because they do not think that it is necessary and, consequently, not provide teachers with the material needed to improve. Conversely, if the evaluator provides feedback during a summative evaluation, the teacher may receive mixed signals regarding the consequences of evaluation. To my knowledge, there are no requirements for school districts to brief or otherwise explain to beginning teachers how ADEPT works or their current stage in the process. This means that how ADEPT functions within a school district may vary from district to district. The result could be a de-standardization of a standardized program: inconsistencies within districts (due to a lack of training and information) and between districts.

If ADEPT is considered as merely summative in nature then it is not very efficient at weeding out ineffective teachers, provided that ADEPT actually measures effective teaching. Mr. Lee could have taught a combined six years without passing South Carolina's minimum standards before being permanently removed from the classroom. But Mr. Lee's termination may have been a result of a failure to develop within the system and not a general lack of teaching ability.

If ADEPT is considered as merely formative in nature then it appears to be designed to give teachers opportunities to *not fail* as opposed to *promoting continual improvement*. This is because the program provides many opportunities to pass formal evaluation and does not emphasize long range improvement until stage four. A teacher's mentor is replaced by another evaluator after the initial year of teaching in an effort to remove bias from the evaluation. However, it appears that in practice this primarily results in a reduced amount of dialogue between teacher and administrator despite the process remaining in a formative stage. This situation is lamentable for both the teacher and the administrators who evaluate them. Teachers do not receive the same quality of direction and mentorship in the months leading up to their high-stakes evaluation. Administrators are tasked with developing relationships with multiple teachers because they must rotate among teachers that they did not mentor in their initial year. Additionally, administrators may be assigned to evaluate teachers in all subject areas at any grade level regardless of their own background. This means that an administrator who was formally educated to teach elementary school may be tasked with providing advice on how to teach Advanced Placement English texts to seniors in high school. The result of this situation is that the ADEPT system expects the administrator to have an area of expertise that may be unfair (and unrealistic). Clearly, the system faces several obstacles in the way of fully achieving its goals of promoting professional development and insuring that the state's public educators are capable.

Research in Education

Educators, being fully invested in the evaluation process, are searching for remedies to ineffective evaluation practices such as the aforementioned issues. They are noting that the traditional method of evaluation, in which an administrator observes a teacher in the classroom, often fails to benefit the teacher because of the way feedback is gathered and then presented (Goldstein & Noguera, 2006). Similarly, Dymoke and Harrison (2006) found that teachers' professional development is primarily rooted in the educational and managerial stance of the school system and not in terms of personal career goals and aspirations. This leads teachers to be evaluated outside of their personal contexts and does not provide motivation for performance. Marshall (2005) makes similar claims in a list of ten reasons why traditional evaluation does not work. Marshall claims that:

1) raters do not evaluate enough teaching,

2) small, individualized lessons do not capture overall performance,

3) raters often evaluate atypical lessons,

4) isolated lessons are an incomplete picture of instructional practices,

5) there is a lack of focus on student learning,

6) the high stakes make it difficult for teachers to learn from evaluations,

7) traditional evaluation makes teachers feel isolated,

8) evaluation instruments limit productive communication,

9) evaluations do not often provide teachers with useful feedback, and

10) raters are often too overwhelmed with other duties to properly supervise and evaluate (p. 728-731).

Clearly, there are a multitude of reasons for dissatisfaction with the current evaluation system.

Because of these concerns, educators are searching for ways, such as peer evaluations or student evaluations (Moore & Kuol, 2005; Beran et. al., 2005), to offset the failings of the traditional process. But there are still concerns regarding the consequential validity and outcomes of one teacher critiquing another's performance because of the impact it might have on the workplace climate if a teacher were to take another teacher's evaluation personally, allowing it to affect his own attitude when conducting evaluations. Additionally, there are many questions surrounding the practicality and usefulness of having students evaluate their own teachers because those students might not understand the motivations behind a teacher's action or take the evaluation seriously. Tucker et. al. (2003) recommends a different solution: teachers keep professional portfolios that are the basis for evaluation. However, this solution is not without its own difficulties. The time and energy it may take to construct a proper portfolio might distract teachers from devoting appropriate time to teaching preparation. It is apparent that the performance evaluation process is a debated and important topic for the education community. The field is currently engaged in discussion about how to make the evaluation process better and is searching for ideas and answers.

Research in Industrial/Organizational Psychology

Industrial/Organizational psychology has examined performance evaluations (often referred to as "performance appraisals" within their discourse community) from multiple facets and provides further insight into the dynamics of evaluation. Specifically, performance appraisal research has addressed areas including environmental and organizational influences, appraisal purpose, evaluation standards, rater goals, and rater accuracy (Jawahar, 2005; Curtis, Harvey, & Ravden, 2005; Scullen, Mount, & Judge, 2003; Murphy et. al, 2004; Sturman, Cheramie, & Cashen, 2005). But despite these various aspects of performance appraisal, there has been very little research regarding the effect of performance appraisal feedback text on rater and ratee. For instance, Davis et. al (2005) and Manshor & Kamalanabhan (2000) acknowledge that more feedback improves performance better than less feedback, but neither article addresses the language used in feedback comments. Similarly, Smither, London, and Reilly (2005) found that improvement in performance is more likely to occur when feedback indicates that change is necessary and that recipients believe change is possible. However, the authors do not examine the particular contents of feedback other than to note it indicated a need for change. *How* the feedback suggests change and the corresponding variance in recipients' receptiveness to the feedback was not evaluated. In other words, current research has neglected to acknowledge the important role of writing and the rhetorical situation in the social construction of the written performance appraisal genre.

I/O psychologists have, however, provided information useful to the present study regarding evaluators. Perhaps most importantly, Murphy et. al. (2004) found that

evaluators give different ratings based upon the perceived goal of the evaluation. This has dramatic implication for evaluation feedback. If an evaluator perceives that the evaluation will only be used for hiring/firing decisions, then she may not consider feedback necessary at all. Even if an evaluator is aware of the goal of the evaluation, he may not provide beneficial feedback because of a failure to incorporate situational influences (characteristics of the situation that may impact performance, such as a student acting out in class) in his evaluation (Jawahar, 2005). The field of I/O psychology has undoubtedly provided us with valuable information regarding feedback and the evaluator, but unfortunately has left the interaction inherent to providing and receiving feedback largely unexplored.

Interestingly, one of the leading evaluation research teams concludes that, because of the counterproductive attitudes and behaviors that may occur as a result of poor evaluation, it is often better to not evaluate at all (Murphy & Cleveland, 1995). But eliminating evaluations altogether is no answer to improving performance. It is clear that most organizations, including school systems, are reluctant to relinquish evaluation systems as a diagnostic and developmental tool because of the emphasis placed upon standardized evaluation. This means that continued research should be able to improve and understand evaluation systems despite their questionable ability to better worker performance. Perhaps approaching evaluations from a different perspective will shed insight into how to reduce the negative effects of evaluation and improve the positive effects.

Research in Professional Communication

Professional communication brings a unique lens to the performance evaluation. It reaffirms the evaluation instrument as a specific communication genre, or communication tool, (Rude, 2004) that functions as a diagnostic used to either make hiring/firing decisions or to develop employees' performance. Those functions indicate that the evaluation instrument should be considered a user-centered document despite little previous scholarly acknowledgement of that status. What classifies the evaluation as a user-centered document is that its primary purpose is not to entertain (as in popular literature) but instead to direct an audience how to achieve some goal. In this sense, the developmental evaluation instrument may be considered in the same vein as, say, an instruction manual on how to operate a device. After reading an instruction manual on how to run a vacuum cleaner, the user should be aware of the steps required in order to properly run the device as well as common mistakes to avoid during operation. Ideally, the evaluation instrument functions likewise, telling users how to perform work activities correctly and providing directions for how to avoid poor performance.

Here, professional communication readily lends itself to reassessment of the evaluation instrument because of its rich history in establishing and evaluating design processes and techniques that insure user-centered documents are user-friendly (which means that the document is easy for individuals to access and use and accomplishes its intended purpose). Particularly, Carliner (2003) offers a three-pronged approach to evaluating usability that is readily applicable to performance evaluation systems. Carliner develops three core value systems that underlie a document's effectiveness: 1)

characteristics based (does the document fulfill a set of criteria?), 2) task-based (does the document enhance the ability of the user to accomplish a task?), and 3) results-based (what value does the document add to the company/organization?) (83). Carliner goes on to argue that the most effective documents incorporate all three value systems. For evaluations in education, Carliner's guidelines for usability mean that evaluation instruments: 1) include the right material in order to fully assess performance (characteristics based), 2) are easy for the evaluator to complete and the reader to understand (task-based), and 3) accomplish the goals of developing better teachers for the school system (results-based). It is likely that such an assessment, which triangulates results from each of the value systems, will highlight ways in which to improve current evaluation systems.

Professional communication also provides solutions for initial document design. As Kain (2005) argues, genres best suit their purpose when all members of a genre's audience participate in its formation. Additionally, Gueldenzoph and May (2002) claim that the best practices for evaluation include not only collaboration during the genre creation stage, but throughout the remainder of the process (implementing feedback and self-assessment of the evaluation system). Jarrett (2005) echoes this sentiment, and instructs that document design should consist of testing several prototypes where representatives of all stakeholders associated with the document have the opportunity to voice their concerns and suggestions for design changes. This finding mirrors that of researchers in the field of education who argue for creating evaluation systems as a collaborative process between evaluators and evaluatees.

It is also important to recognize that the performance evaluation is socially constructed. Since the time of the sophists, rhetors like Gorgias have argued that truth is not absolute and is a result of a rhetor and an audience agreeing upon meaning. More recently, studies have shown truth to be socially constructed in a variety of knowledge-making fields. For example, Latour and Woolgar (1986) found that scientists, who were traditionally thought to view the world has having Absolute Truth, construct scientific fact socially. Scientists' knowledge claims were codified by Latour and Woolgar and they found that ideas generally accepted by the scientific community did not need qualifying, but those that were more tentative used modifiers such as "likely" and "probably" and invoked other researchers. Other studies have found similar results in scientific writing (Myers, 1985).

In other words, scientists had to argue for their "truths." Burke (2001) further demonstrates that truth is socially constructed with the concept of "terministic screens." Terministic screens apply to all language and indicate that, at the same time meaning is being made, meaning is also denounced. These terministic screens differ from discourse community to discourse community and on an individual basis. For example, the word "good" carries multiple meanings. For one individual "good" may indicate "not bad," while another individual infers that "good" means "pretty well." The key difference in these definitions is that the first is phrased negatively whereas the second is positive. In the case of developmental evaluations, meaning is constructed when the evaluator and evaluatee agree on an assessment of performance and the steps necessary to improve that performance. For administrative evaluations (those used to make hiring/firing decisions),

meaning is made when the supervisor tasked with making those decisions accepts performance evaluation data as a truthful indicator of ability. As one might infer, the difficulty in making meaning might occur when the evaluator says "good" meaning "pretty well" and the evaluatee infers "good" as "not bad." It's the difference between a performer thinking she has had either above average or average performance.

The performance evaluation also creates a rhetorical situation (Bitzer, 1968). As Bitzer explains, every rhetorical situation consists of three components: 1) an exigence, or an important reason for communicating; 2) an audience that is affected and interested by the discourse; and 3) a set of constraints that shape and modify the exigence. In this case, the rhetorical situation is between the evaluator and evaluatee with the evaluation itself as the exigence. Goodall, Wilson, and Waagen (1986) explain the rhetorical situation in the context of a face-to-face evaluation interview. The two primary exigencies for the evaluator are: 1) to convince the evaluatee of the accurateness of the performance assessment, and 2) to create a dialogue to negotiate future goals and changes in performance. The primary exigency for the evaluatee is to accept or modify the assertions regarding their performance made in the evaluation. This is done through engaging the evaluator in dialogue. However, this rhetorical situation is altered dramatically when the evaluatee does not have the opportunity to engage in discussion with the evaluator outside the evaluation instrument and the included feedback. When the evaluatee only receives written feedback from the evaluator, the stakes are much higher. The evaluator must make a case for the assessment and trust that the evaluatee comprehends the evaluation and any suggestions for performance improvement. This is

the particular rhetorical situation that was first illustrated in the vignette that opened the chapter. Despite an evaluator's (Administrator Johnson) and evaluatee's (Mr. Lee) best efforts at accomplishing the evaluation task, the evaluation instrument and system held much of the power within the rhetorical situation. This led to both parties being dissatisfied with the evaluation's outcomes.

In summary, the field of professional communication approaches evaluations from a theoretical perspective different from the areas of education and I/O psychology. Namely, the focus is shifted to the interaction between evaluator-instrument-evaluatee and how performance "truths" are created within the rhetorical situation. This approach should provide insight into the evaluation process of teachers that educators might not be aware of and that I/O psychologists have missed, despite their wealth of knowledge.

The present study is intended to examine the performance evaluations of teachers from the professional communication perspective. This is accomplished by evaluating the impact that language may have in written performance feedback. In the next chapter, the methodology of this study is explained and its limitations are discussed.

CHAPTER TWO

METHODOLOGY

This study involves the testing of preliminary research regarding effective communication in performance appraisal comments which suggested that there are ways raters can improve the effectiveness of written feedback. Specifically, raters can improve evaluations by making comments that reflect proper use of content, word usage, and structural order. These preliminary conclusions are based on analysis of in-depth questionnaires completed by three student teachers in March 2006 in a small pilot study (see Appendix B). The student teachers were all participating in the secondary education program at a respected southeastern private university. They had just finished their term of student teaching, which includes several evaluations from raters who work within the school district as well as university professors. Each student teacher completed the questionnaire (see Appendix B) and submitted the completed version electronically. The questionnaire was designed to solicit responses concerning effective communication in multiple ways. Participants responded to both actual and ideal situations and were asked to reflect upon their personal experiences with how communication in performance evaluations succeeded or failed. An open-coding analysis of the questionnaire data found nine specific aspects that impact evaluation comment effectiveness. These findings provided the basis for the present study, which seeks to collect and analyze information regarding the effective use of language in performance evaluations.

Current Sample

School districts were contacted via email after receiving permission from the South Carolina State Board of Education (see Appendix C). One school district that covers a large metropolitan area in a southeastern state participated in the study. The school district was chosen on the basis of size in order to obtain a sample in excess of 1,000 possible participants. This was a purposeful sample since the targeted demographic included both teachers and the administrators who evaluate them. Organizational restrictions required that the survey invitation (see Appendix D) was forwarded to possible participants by the district's research coordinator. Researchers were not allowed to obtain specific contact lists or email listservs, but were assured that the survey was sent to all teachers and administrators within the district that were involved in the appraisal process. The target response rate was thirty percent; however, the actual response rate was much lower at two percent. There was a total of 26 responses, two of which were administrators who performed evaluations. Mortality during the survey process further reduced this number to 17. However, both administrators completed the survey, which allowed for tentative comparisons of the two groups (evaluators and evaluatees).

The school district utilizes the state's system for Assisting, Developing, and Evaluating Professional Teaching (for more information on the ADEPT system, see Appendix E). The participants in the pilot study were evaluated under the ADEPT system and were student teachers in the same school district. This means that the same evaluation practices found in the pilot study were replicated in the present study.

Part of the rationale for selecting a larger school district was to gather information from a diverse participant pool. This study had a moderately diverse participant pool despite the low number of completed surveys. Thirteen of the 15 teachers responded to the subject area taught item, of which seven were elementary, two were secondary school English, two were secondary school mathematics, one taught secondary school social studies, and one taught special education. Sixteen participants responded to the gender item and thirteen participants were female. One administrator was male and the other was female. Most participants (11 of 17 responding) had been employed in their current position 1-5 years, three 5-10 years, one 10-15 years. Both administrators had been employed in their current position 1-5 years.

Survey Design and Measurement

In general, the purpose of the survey (see Appendix F) is to collect information regarding the effective use of language in performance evaluations. Specifically, the survey was designed to evaluate the three categories identified in the pilot study that influence evaluation comment effectiveness: word usage, word content, and structural order.

*Category 1: Word usage***.** The pilot study found that word usage influences the effectiveness of communication in performance evaluations. A balanced tone that distances the action from the actor in cases of negative behavior and that avoids speaking in absolutes constitutes effective word usage.

This section of the survey evaluated word usage by presenting samples of criticism that vary in tone, use of passive voice, and inclusion of "absolute" words such as "never," "always," and "failed." Each sample comment was designed to be as similar as possible to others, except for the inclusion of the words and phrases that alter tone. The samples were evaluated on a Likert scale that rates potential to produce positive developmental change (improved work performance). Additionally, participants were provided with a free response area to list any additional words or phrases that are perceived as counterproductive to producing effective feedback. Specifically, this section evaluates hypothesis one, which stated that *teachers are more likely to identify ineffective word usage in evaluations than administrators.*

*Category 2: Content***.** Content was also identified in the pilot study as an area critical to effective communication in performance evaluations. Effective content demonstrates that the rater is conscious of the evaluation purpose, situational context, and explicitly qualifies evaluative comments.

The survey evaluated this category by presenting a performance observation scenario similar to the examples provided in the pilot study data. This scenario is followed by a list of possible remarks made concerning that scenario. The remarks were constructed so that they have varying degrees of contextual awareness and some remarks were made without qualifiers. Participants choose what they consider to be the option most likely to produce positive developmental results from the list of remarks. However, none of the remarks represent the hypothetical, best possible feedback. This is because participants were provided with a free response area to write their own feedback

comment regarding the scenario. The rationale for the design of this section was to encourage participants to think of feedback for themselves, and it should provide qualitative data that demonstrates the effective use of content. Specifically, this section evaluates hypothesis two, which stated that *there is a strong positive correlation between contextual understanding and effective language in performance evaluation remarks. However, administrators and teachers do not share the same level of contextual understanding, leading to discrepancies in the perceived effectiveness of feedback.*

Category 3: Structural order. There was some indication in the pilot study that structural order also affects the effectiveness of written communication in performance evaluation. Pilot study data indicated that negative statements should be prefaced by positive ones, and negative statements should not be in the form of a question.

To evaluate structural order, the survey included a section of sample written feedback, as it might appear on an evaluation form. Each example of feedback was evaluated on a Likert scale that rates potential to produce positive developmental results. The samples of feedback were designed to reflect the pilot study's examples of general effective and ineffective structural order. For instance, one sample includes a negative comment prefaced by a positive comment, and another posed a negative comment in question form. Specifically, this section evaluates hypothesis three**,** which stated that *neither administrators (raters) nor teachers (ratees) will consider negative statements in question form as effective communication.*

Data Collection

Participants completed an online survey (see Appendix F) hosted through surveymonkey.com. The method of delivery (email invitation) is probably partially responsible for the low response rate. As both MacNealy (1999) and Rogelberg et. al. (2002) explain, one of the disadvantages of computer-delivered surveys is that people are now inundated and often disregard emails. However, this method of delivery was the most feasible, as the organization communicated frequently via email and all potential participants had ready access to email and the web. In order to control factors threatening validity, the survey was open to all participants for the same amount of time and each participant completed the same survey. The survey was available online from March 31-April 18, 2007. The length of data collection, while long for an online survey, was done at the request of the participating school district in order to allow teachers and administrators ample time and opportunity to participate in the research. A "one survey per person" limitation using online cookies was imposed in order to prevent the same computer from submitting multiple responses. Order bias was avoided by randomizing the questions within each survey section for each participant. The survey was taken at the time and place of the participants' choice. Preliminary evaluation of the instrument utilizing two test subjects found that the survey took approximately 10 minutes to complete.

Procedure

All data, including demographic material and survey responses, were evaluated using surveymonkey.com's analysis function and exported into a MICROSOFT EXCEL spreadsheet for verification. Analysis of data included identifying correlations and discrepancies across a broad spectrum of criteria. Each category (content, word usage, and structural order) was evaluated in order to determine the accuracy of the pilot study findings and for verification of the hypotheses. Additionally, a number of other, external characteristics that may affect the perceived effectiveness of language in performance evaluations were analyzed to determine significance.

The content category was analyzed to determine if there was consensus between the preliminary results regarding effective use of content in evaluation remarks and the survey responses. The word usage category also was analyzed to determine if there is agreement regarding how tone, passive voice, and particular words affect the perceived effectiveness of evaluation comments. Finally, the structural order category was analyzed to determine if there is a general consensus regarding the way comments should be structured within evaluations. Demographic trends and variances that might occur between administrators and teachers (such as subject area taught and length of employment) that may affect perception of how content, word usage, and structural order influences effective communication in evaluations were also considered.

Through measuring participants' responses to each of these three areas the study intends to not only test the hypotheses that correspond to each category but also to illuminate features of the rhetorical situation of the performance appraisal. While direct

observation of the evaluation process was not an option, it was hopeful that the structure of the scenario situated participants as either the rhetor or audience (depending on the survey item) in the rhetorical situation. However, given the methodology, it is important to understand some of the issues regarding survey research that may impact the analysis of data and subsequent conclusions.

Limitations

MacNealy (1999) discusses several threats regarding validity when conducting empirical research in professional communication. Given the inherent limitations of the present study, it is necessary to discuss some of these issues in order to better understand its methodology and resulting data.

Internal Validity

History. Participants may have had experiences prior to taking the survey that impact the data. For example, a teacher may have recently had a poor evaluation that caused him to have a distaste for evaluations, no matter the level or type of feedback. However, since the majority of the participants had been working in their current position for the same amount of time, it is reasonable to expect that they have been through a similar number of evaluations. Since this study is designed to improve the feedback in evaluations, it is not necessarily a threat to validity if a participant had a "bad" or "good" experience with evaluation feedback. Instead, such participants might provide valuable insight into the mechanics of feedback.

Mortality. Nine participants did agree to take the survey only to stop answering survey items. It is possible that these participants dropped out for the same reason. This group of subjects may have had a particularly bad experience with feedback that made the survey traumatic, or they may have not been affected by evaluation feedback enough to consider completing the survey worthwhile. However, there are multiple other reasons for a participant to drop out: lack of time, unforeseen interruptions, internet connectivity issues and so forth. Even if these participants failed to answer survey items because of an evaluation-related reason, more than half of the participants *did* complete the survey. This indicates that the research is relevant and worthwhile.

Diffusion. It is impossible to determine if any survey participants came into contact with each other and discussed the survey with a participant prior to completing the survey. Such contact prior to participation may have affected the way an individual answered survey questions. However, it is reasonable to assume that diffusion was not likely for several reasons: 1) there was a small number of participants, 2) there was some distance between each response during collection, and 3) the large number of schools included in the district.

External Validity

Bias Error. There is some possibility that self-selection might have occurred. There may have been a particular personality trait shared by the subjects that caused them to participate. Participants were informed that this study was part of a graduate thesis, so they may have been encouraged to complete the survey in order to assist the researcher's education. However, other demographic factors that often lead to bias error do not

appear to present. The sample participants earned roughly the same salary (public education pay scales are standardized), were representative of the gender makeup of the organization, and were probably of similar age (according to years employed).

Generalizability. There are several factors to consider when assessing the generalizability of the present study. Primarily, it is acknowledged that the low response rate is not statistically significant. This means that there were not enough responses in order to scientifically claim that a hypothesis was confirmed, partially supported, or not supported. However, it was possible to assess the data gathered when evaluated in conjunction with the theoretical framework regarding the nature of performance evaluations that has been established. This framework informs the analysis of data and the subsequent discussion of conclusions from the data.

Performance evaluations in other contexts and occupations may or may not utilize individualized written feedback as part of the evaluation process. Many organizations rely upon performance appraisal instruments that do not require feedback or use stock phrases, such as "try other strategies," making evaluations of teaching unique. There are several reasons for not including feedback in an evaluation: 1) it takes more time to provide feedback, 2) the rater may be unable to provide appropriate feedback, and 3) there is sometimes little variance in the feedback that may be required in occupations where actions are simple and repetitive. However, in any occupation where performance is much more complex, feedback may be useful for improving performance. It is hopeful that this research will, in part, encourage organizations to utilize feedback effectively.

Thus, while the results of this study may not generalize to all current performance appraisal settings, it does suggest the usefulness of such feedback.

We may, however, tentatively generalize the study to the extent of the school district that it represents. This alone satisfies the utility of the research: the school district is large, is representative of other public school districts in its practices, and may benefit greatly from the study's findings. Within this context, the participant sample is moderately representative of the organization: primarily female, employed between one to ten years, and there was a low ratio of evaluators to evaluatees. These factors indicate that the study sample, while small, is appropriate to the research.

Construct Validity

Methodological choices were governed by permissibility and time constraints. It is acknowledged that triangulation of the data would have been favorable, but it was not feasible. The present study is designed to test the reliability of the pilot study's findings, which were gathered through detailed, interview-quality questionnaires. If the present study's findings confirm the preliminary results of the pilot study, then we may begin to build construct validity. This is because data was first gathered through in-depth questionnaires, then assessed using multiple choice responses and open-ended essay responses. Future studies should be able to build upon the findings. This study should be repeated on a larger scale, and we should also obtain further qualitative data from evaluators and evaluatees. Additionally, obtaining copies of actual written evaluations

would allow discourse analyses of feedback. Obtaining such data is unlikely for the outside researcher, but it could confirm construct validity.

Summary

A survey testing the findings of the pilot study is the first step in confirming what constitutes effective feedback in performance evaluations. All available measures were taken to preserve the validity of the findings. There are several factors acknowledged that threaten validity (namely, the small sample size), but the demographic variety of the sample alleviates many of those concerns. In the next chapter, we will discuss the findings of the survey. However, the results should be interpreted with some caution given the methodological limitations.

CHAPTER THREE

RESULTS

The following is a dissemination of the results of the survey. All items for each of the three categories (word usage, content, and structural order) are examined. All tables identify the corresponding survey item that was evaluated by participants. Significant findings regarding each hypothesis and a review of demographic differences that may impact the data are reported. This data provides implications regarding the understanding of the rhetorical situation of performance evaluations and is discussed in Chapter Four.

Word Usage

In Tables 1-4, the number of responses to each answer for each item of the *word usage* category are reported. The survey results support pilot study findings regarding word usage. In Table 1, the majority of respondents identified the feedback sample, "you failed to manage the classroom properly because you did not maintain authority," as one that could adversely affect performance. This item included the word "failed" which was identified in the pilot study as a word that should not be used. It is likely that some respondents felt that this statement might improve performance because it does provide a specific regarding how classroom management was not successful. The two administrators included in this sample did not concur, with one selecting "may produce improvement" and the other choosing "may adversely affect performance."

Table 1
"You failed to manage the classroom properly because you did not maintain authority."

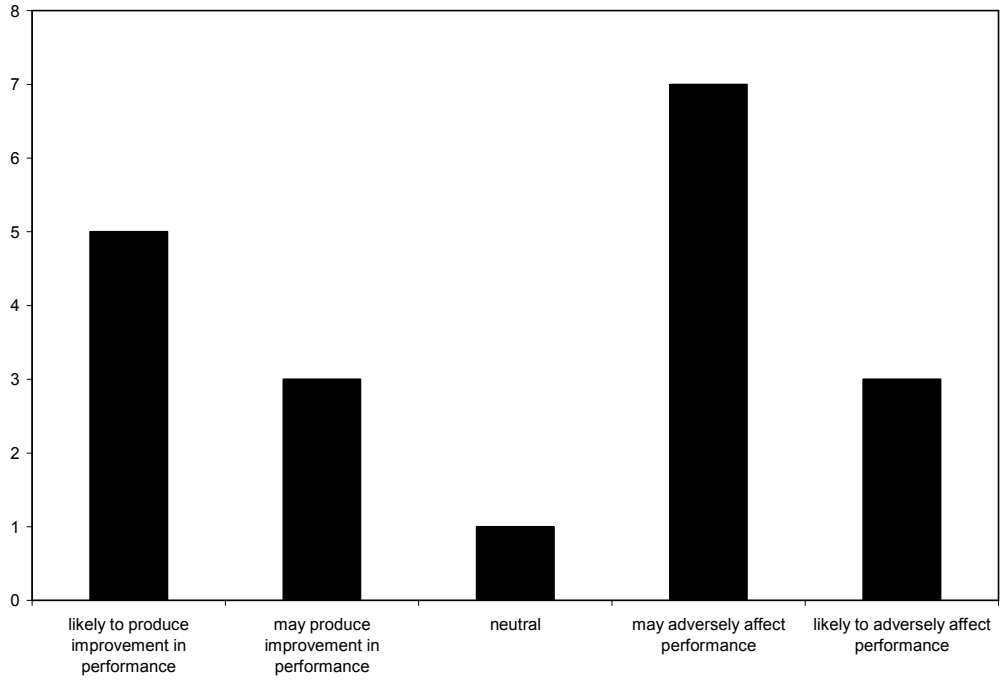
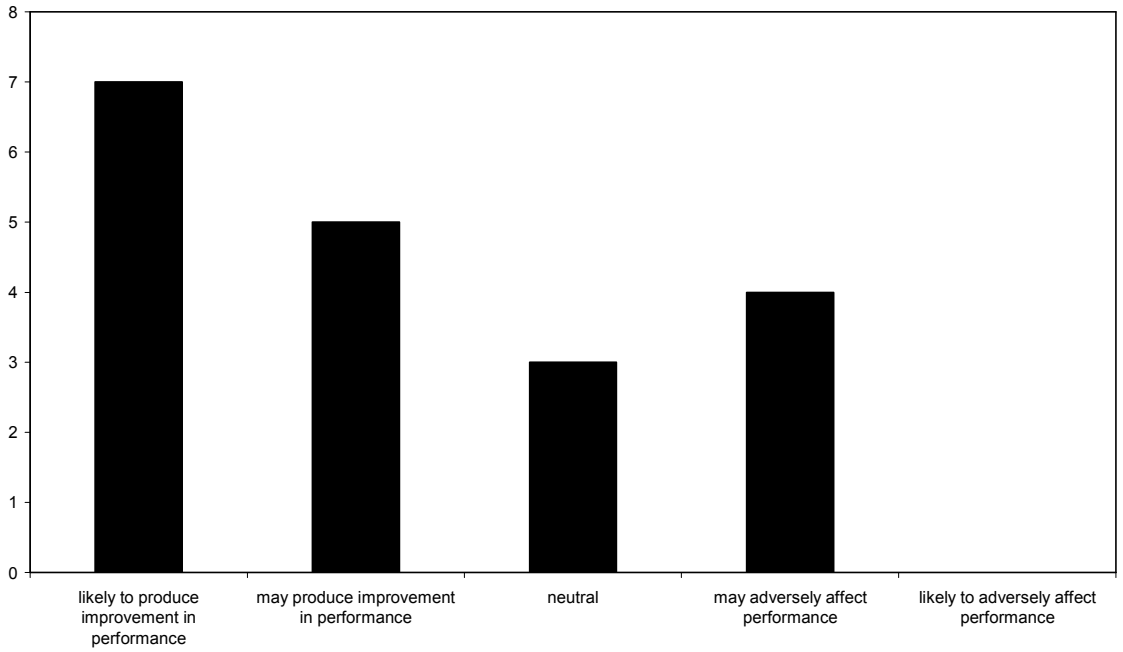


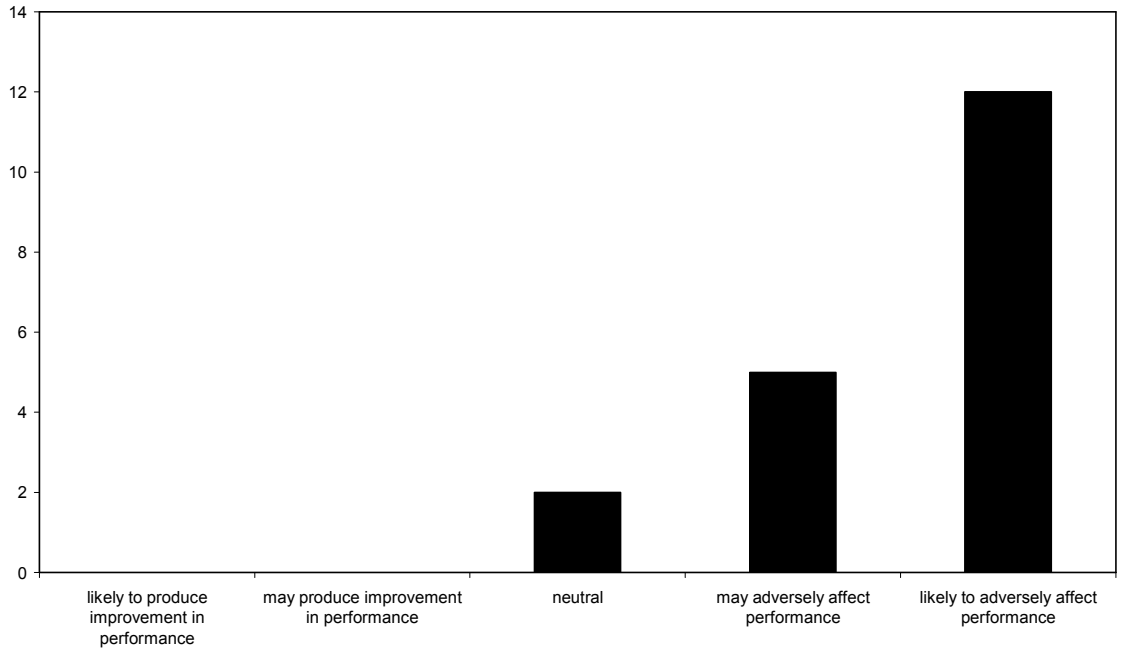
Table 2 indicates that participants felt that the second feedback statement, "you might improve your classroom management by maintaining authority," was appropriate feedback, though seven respondents indicated that the statement may cause no change or adversely affect performance. These results were anticipated by the pilot study, as the statement has no words that carry negative connotation. However, the statement did not distance the actor from the action (utilizing the pronoun "you"), which was identified as counterproductive word usage. This may account for the seven respondents that indicated the phrase may adversely affect or have no affect on performance. The two administrators included in this sample did not concur, with one choosing "neutral" and the other selecting "may produce improvement."

Table 2
"You might improve your classroom management by maintaining authority."



The results displayed in Table 3 strongly indicate that the feedback phrase, "because you did not maintain authority I have never seen poorer classroom management," was not appropriate for improving performance. These results were expected as the feedback phrase used included two words (never and poorer) identified as counterproductive to improving performance and did not distance the actor from the action. The two administrators included in this sample concurred that the phrase would adversely affect performance.

Table 3
"Because you did not maintain authority I have never seen poorer classroom management."



The feedback phrase in Table 4, "one way to improve classroom management would be to maintain authority," exemplified appropriate word usage for improving performance. Participants identified this phrase as the most likely to improve performance and no respondents indicated that the phrase may harm performance. The two administrators included in this sample both indicated that the phrase may produce improvement in performance.

Table 4
"One way to improve classroom management would be to maintain authority."

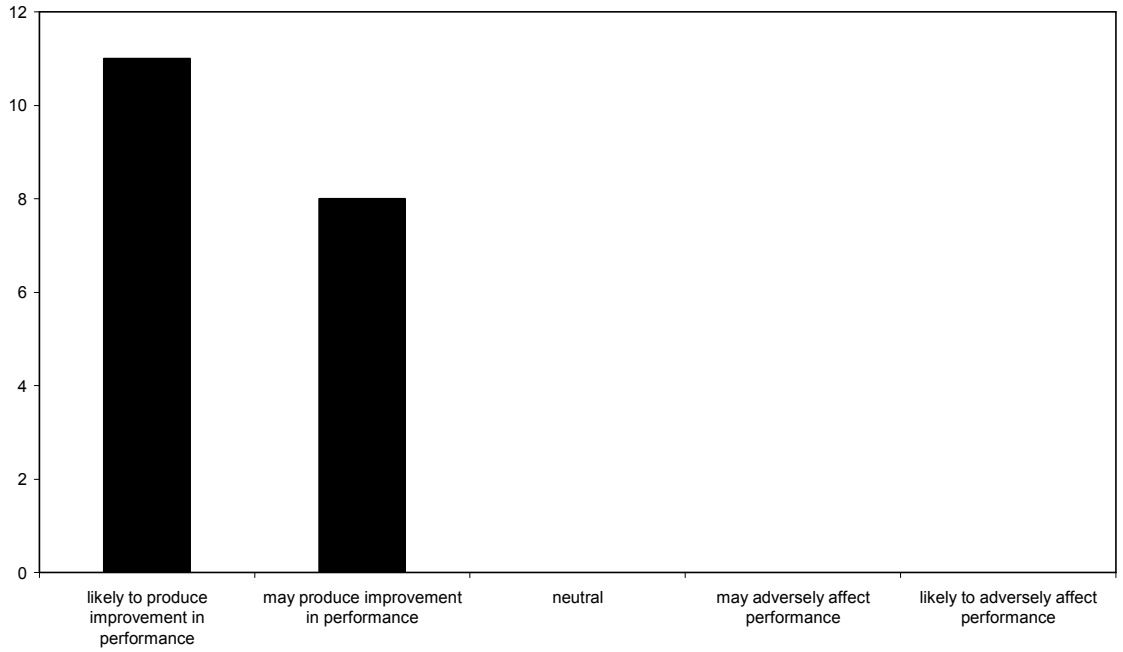


Figure 1 presents significant findings from item six on the survey, which asked participants to list counterproductive words and phrases often found in performance evaluations. The words and phrases identified, especially those that carry negative connotations or are absolutes, are consistent with those of the pilot study.

Figure 1
Words Considered Counterproductive to Improvement

-No other alternatives to fix the issue

-Not you; you failed

-Poor; poorer

-*Absolutely no; teacher obviously not**

-Never; failed; might improve

-*always; required*

-Worst I have ever seen; I was appalled

*Items in italics indicate administrator responses

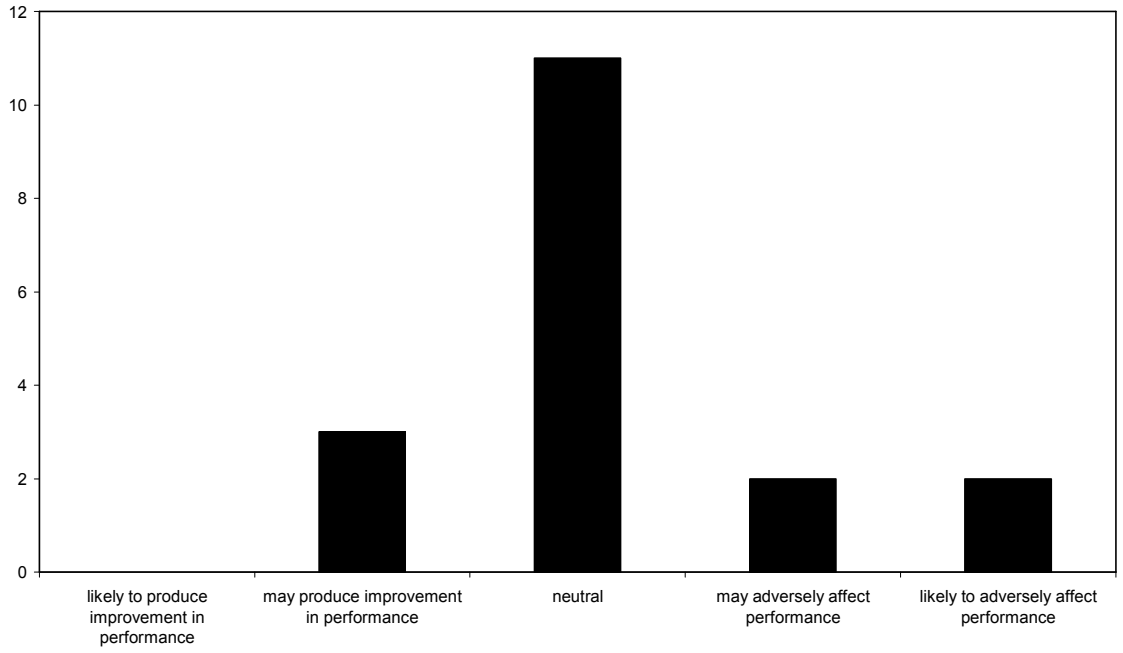
Tables 1-4 strongly suggest that there are instances of word usage that may adversely affect performance. Additionally, participants identified similar counterproductive words and phrases for the open response question. The two administrators in the sample were as able to identify inappropriate word usage as the teachers. Thus, the results indicate that hypothesis one, which stated that teachers are more likely to identify ineffective word usage in evaluations than administrators, may not be supported. However, the results do indicate that words that imply no room for change and absolutes are counterproductive to encouraging positive change in performance.

Content

Next, in Tables 5-8 the number of responses to each answer for each item of the *content* category is reported. The participants rated feedback phrases based on a short

performance scenario. The results in Table 5 strongly indicate that this phrase, "try other methods to achieve classroom management," would produce no change in performance. Both administrators included in this sample concurred with the majority, believing that this statement would cause no change. This phrase included no qualifiers, and indicates the rater is marginally aware of the instructional situation. It would not exemplify effective feedback.

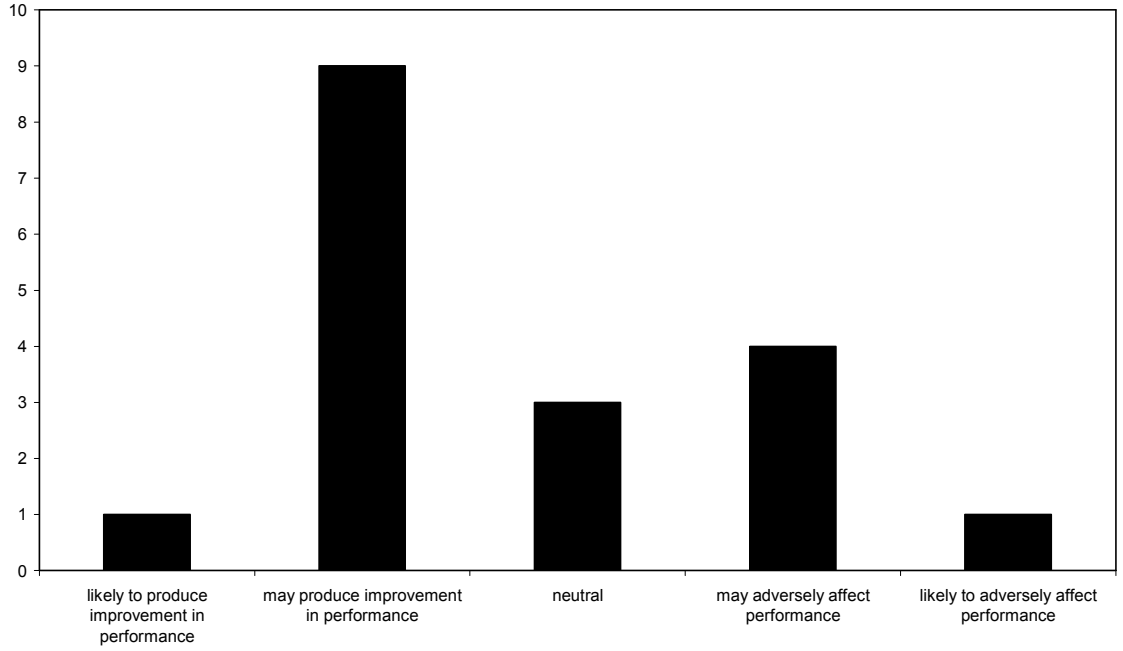
Table 5
"Try other methods to achieve classroom management."



The results in Table 6 indicate that this phrase, "while you were doing your lesson plan what were your accommodations," might improve performance. The feedback phrase indicates an awareness of the instructional situation and hints that the teacher should have accommodations (plans prepared that help teachers responded appropriately to students with special behavioral and learning needs) to deal with particularly talkative

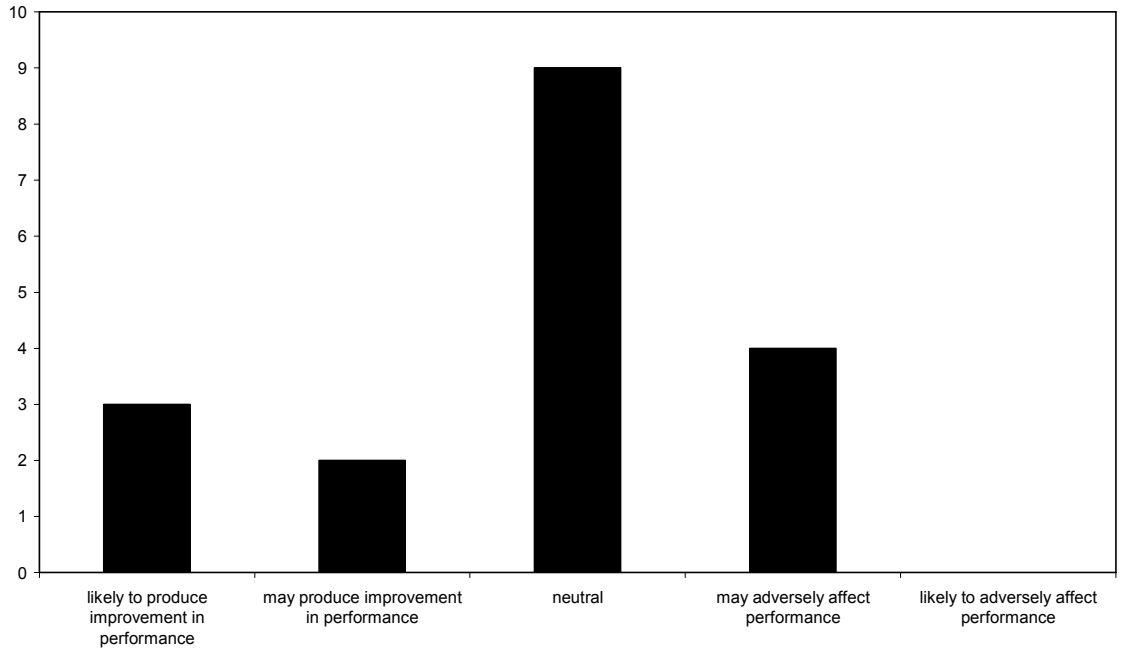
students. However, this phrase does not represent ideal effective feedback because it is somewhat ambiguous. The uncertainty of the feedback is probably responsible for the variety of other responses to this item. The two administrators included in this sample differed, with one selecting "may produce improvement" and the other selecting "neutral."

Table 6
"While you were doing your lesson plan what were your accommodations?"



The results in Table 7 suggest that this phrase, "good way to quiet down your kids," would produce no change in performance. The two administrators included in this sample chose differently. One selected "neutral" while the other believed it may adversely affect performance. This phrase included no qualifiers, and is marginally aware of the instructional situation. It would not exemplify effective feedback.

Table 7
"Good way to quiet down your kids!"



The results in Table 8, which assessed the feedback phrase "nice way to offer a review of material," show that participants were divided regarding whether or not the statement would improve or produce no change in performance. This demonstrates that while the phrase was positively framed and showed contextual understanding, there was still some ambiguity regarding the intent of the phrase. Recalling the evaluation scenario (the teacher used worksheets to calm students after a Friday pep rally), this phrase might indicate falsely that worksheets are always the answer to talkative students. It would not exemplify effective feedback. Both administrators included in this sample marked that the phrase would produce no change in performance.

Table 8
"Nice way to offer a review of material."

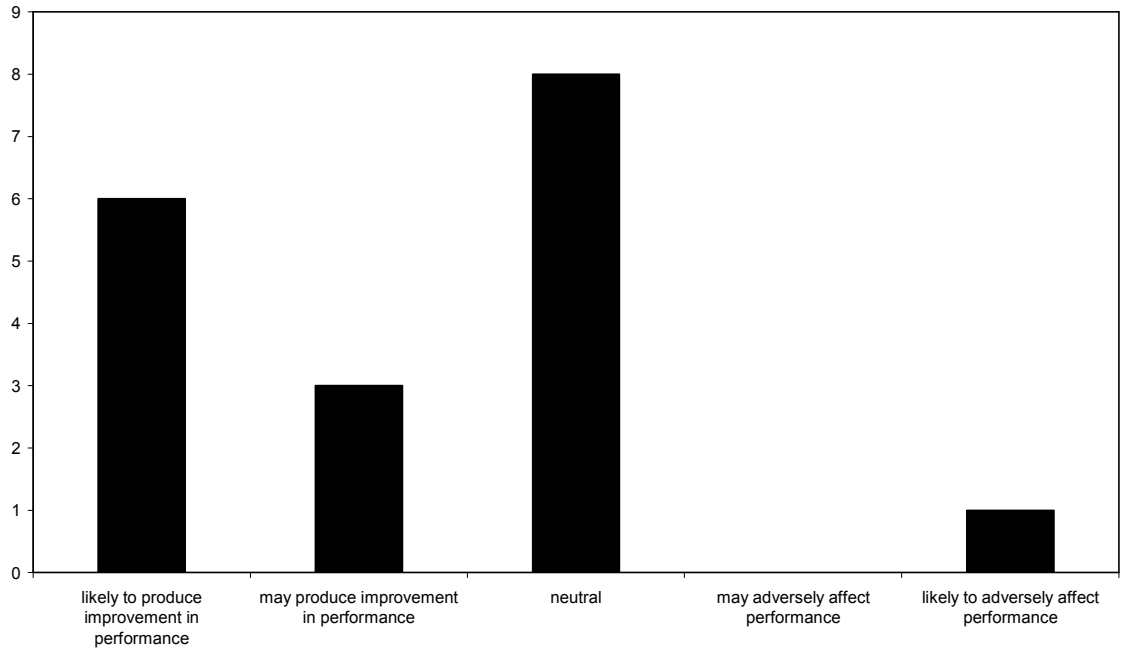


Figure 2 presents significant findings from item 11, which asked participants to write a feedback response to a performance scenario in which a teacher used worksheets to gain order in the classroom after a Friday pep rally. Some of the responses are much more effective feedback. Responses such as the third example demonstrate that it is possible to write feedback content with a high level of situational awareness with minimal information. The longer responses tend to be the most complete, and while the remark by the administrator is superior to the Likert rated examples, it does not emphasize an understanding of what is usually an instructional situation that is difficult to maintain.

Figure 2
Open Response to Performance Scenarios

-How did the worksheet tie in with the pep rally?

-This was an effective way to settle your class after an assembly, but how can you make sure that the students who struggle without direct instruction will remain on task?

- I can understand the need to maintain control on Fridays. The end of the week is a tough time to work with students. However, what are some other ways you can, in addition to the worksheets, improve upon classroom management? What are some ways you could accommodate for those students who might need more verbal or kinesthetic methods?

- *It is obvious that you know how to manage class behavior but do not lose effective instructional practice in the process.*

- I would wonder how a review sheet to be turned in at the end of class on a Friday would be helpful in the following week.

- tough day to come observe with all that is going on - good job on controlling the situation

- Having been in your shoes, I think she did the best she could do under the circumstances -- pep rally and Friday. The kids were excited and happy. She didn't try to boss or bully them into doing something, she used their energy to review material with worksheets.

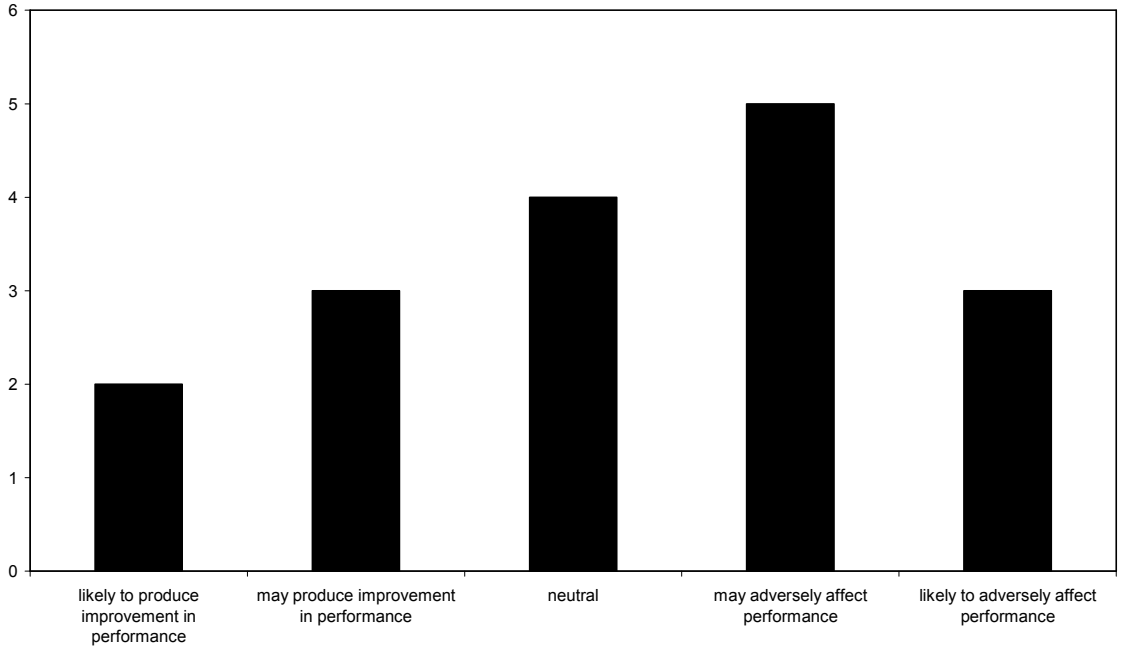
*Items in italics indicate administrator responses

Hypothesis two stated that there is a strong positive correlation between contextual understanding and effective language in performance evaluation remarks; however, administrators and teachers do not share the same level of contextual understanding, leading to discrepancies in the perceived effectiveness of feedback. Results indicate that part one of hypothesis two may have been confirmed if statistical significance had been reached: feedback that demonstrated an awareness of the observation context was preferred in the open responses while the example phrases received mixed ratings. However, administrator responses to the content section of the survey do not provide enough information in order to assess part two of hypothesis two.

While the administrator open response is not as effective as some of the teacher responses, the difference in responses to the Likert items show that part two of the hypothesis is inconclusive. Thus hypothesis two was partially supported.

Structural Order

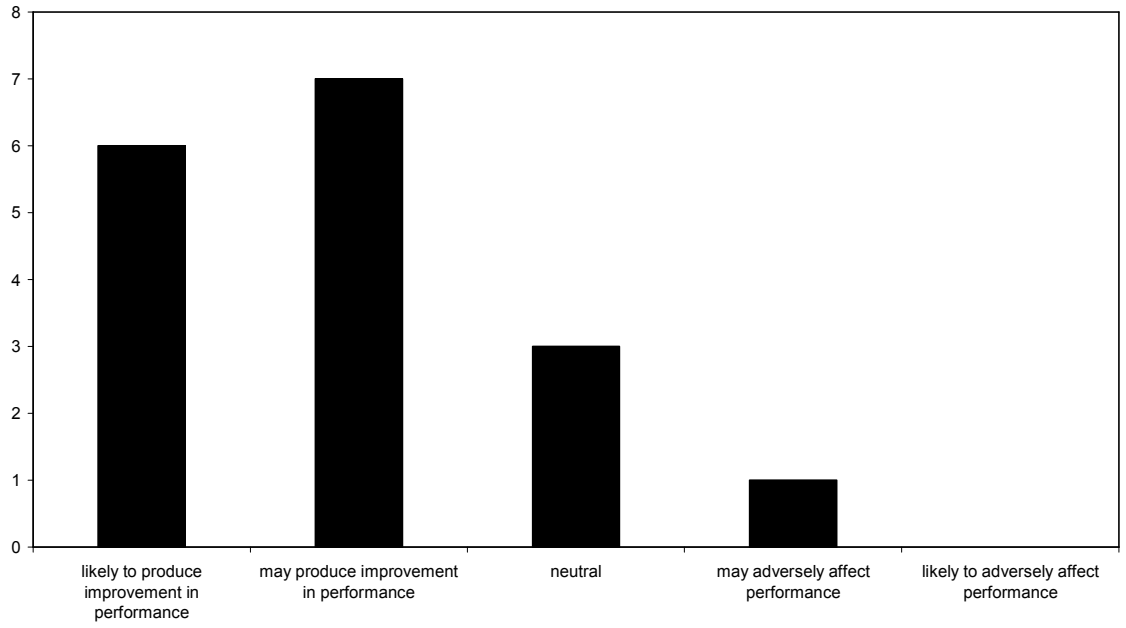
Table 9
"Why did you discipline student A that way?"



In Tables 9-11, the percentage of responses to each answer for each item of the *structural order* category are reported. The results in Table 9 suggest that this phrase, "why did you discipline student A that way," would not affect or adversely affect performance. Both administrators included in this sample marked that this phrase may adversely affect performance. These results are concurrent with the pilot study findings, which suggest that questions that are negatively framed or accusatory would not be considered effective performance feedback.

The results in Table 10 strongly indicate that this phrase, "you handled your class well. Perhaps next time you will consider disciplining student A by…," is effective feedback. It follows the structural order determined in the pilot study to be most likely to improve performance: it contains a positive affirmation followed by a suggestion for change. The two administrators included in this sample concurred with the majority of respondents, marking that the phrase may produce improvement in performance.

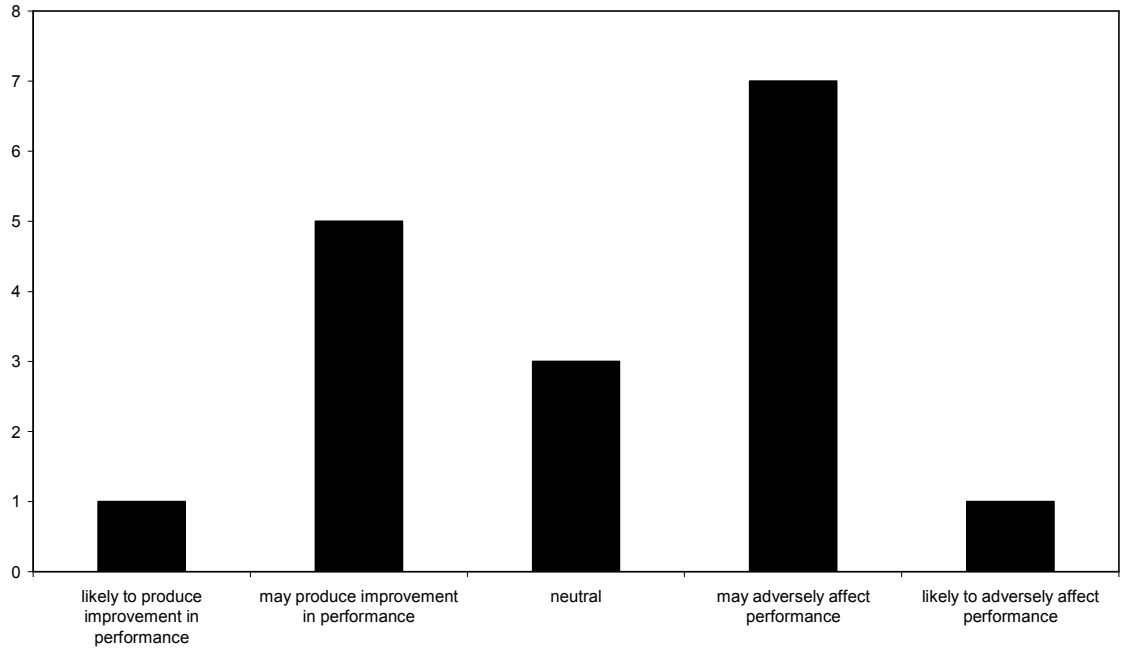
Table 10
"You handled your class well. Perhaps next time you will consider disciplining student A by…."



The results in Table 11 indicate that this phrase, "you did not discipline student A appropriately. But you handled the class well," is ambiguous, and that it may harm, improve, or cause no change in performance. These results are as expected, because the phrase mixes ineffective and effective structural order. It does not contain a question, but it prefaces a positive comment with a negative. The two administrators included in this

sample differed in their responses, with one believing the phrase may produce improvement and the other selecting that it may adversely affect performance.

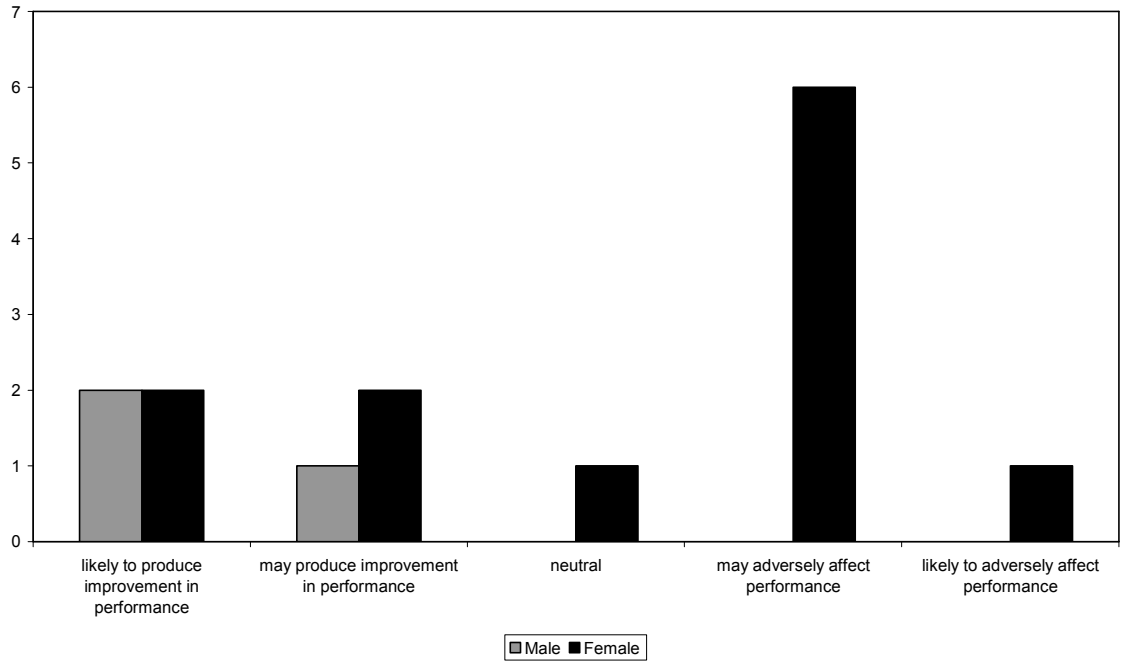
Table 11
"You did not discipline student A appropriately. But you handled the class well."



Thus, the results indicate that hypothesis three, which stated that neither administrators (raters) nor teachers (ratees) will consider negative statements in question form as effective communication, was confirmed. Feedback that contains a positive phrase followed by a suggestion was found to be most likely to improve performance, while the feedback examples that contained some of the components of effective feedback received mixed assessments.

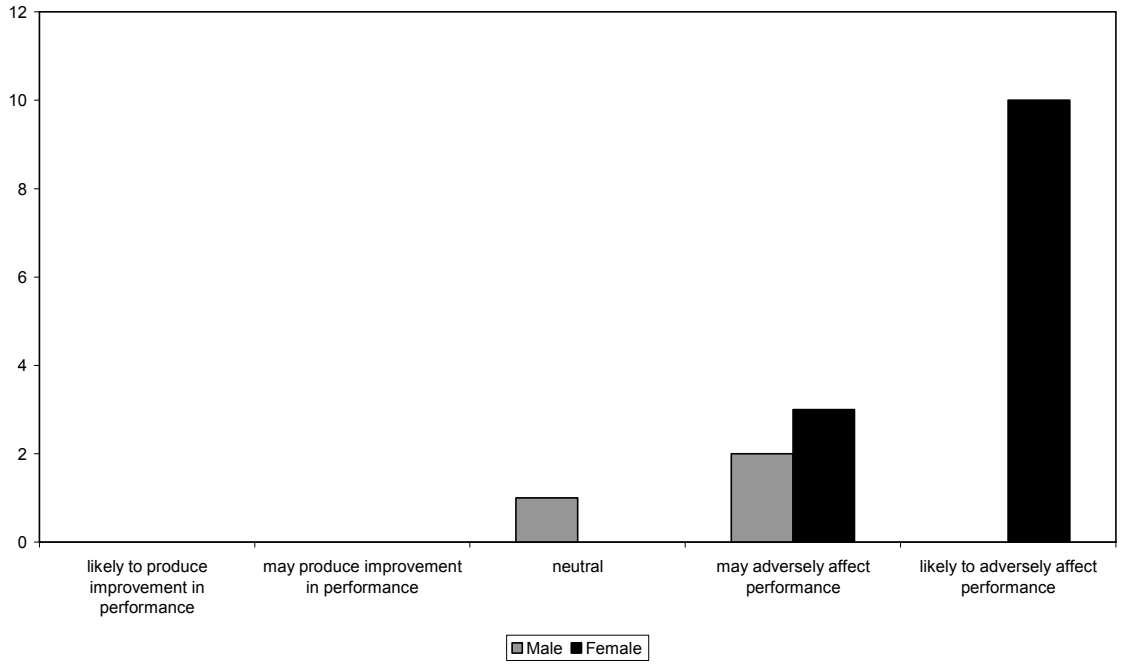
Demographic Considerations

Table 12
"You failed to manage the classroom properly because you did not maintain authority."



Finally, to explore other possible correlations within the demographic data, each completed survey was filtered by subject area (if participant was a teacher), work experience, and gender. Relationships within the data were difficult to assess given the small sample size. However, a few tentative correlations were discovered within the teacher's subject area and gender categories. Teachers of English were more likely than all the other subject areas combined to identify effective and ineffective feedback across all three categories. Male participants did appear to be less adept at identifying appropriate feedback than their female counterparts, but the sample size is too small to consider gender as a contributing factor (see Tables 12-14). Work experience did not appear to be a contributing factor.

Table 13
"You might improve your classroom management by maintaining authority."

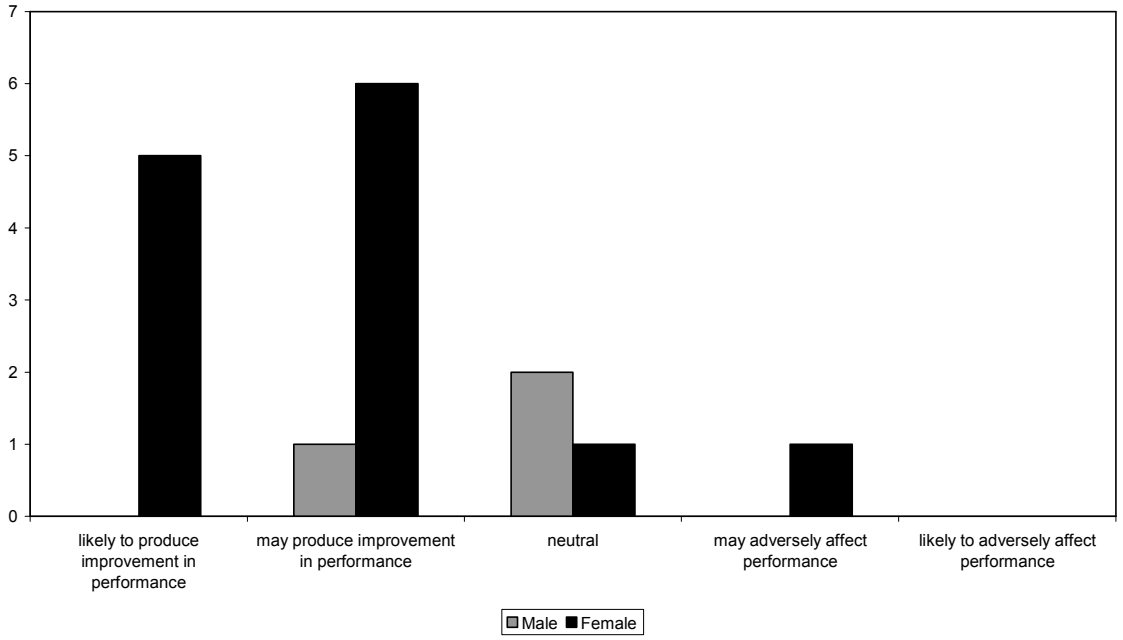


Summary

Overall, results are consistent with the findings of the pilot study when statistical significance is not considered. The characteristics necessary for effective written feedback in the word usage, content, and structural order categories were confirmed through the Likert evaluations of feedback samples. The open responses to the word usage and content categories give additional validity to the findings, as they reflect examples provided in the pilot study. Hypothesis one, which stated that teachers are more likely to identify ineffective word usage in evaluations than administrators, was not supported. Part one of hypothesis two was confirmed: feedback that demonstrated an awareness of the observation context was preferred in the open responses, while the example phrases received mixed ratings. However, the data was insufficient to support

part two of hypothesis two, that teachers (evaluatees) and administrators (evaluators) do not share the same level of contextual understanding. Hypothesis three, that neither administrators nor teachers will consider negative statements in question form as effective communication, was supported. Additionally, results indicate that females and teachers of English may be more likely to identify effective and ineffective feedback. In the next chapter, we will further discuss these results, the implications that they carry, and directions for further research in performance evaluation feedback.

Table 14
"You handled your class well. Perhaps next time you will consider disciplining student A by…."



CHAPTER FOUR

CONCLUSIONS AND FUTURE DIRECTIONS

To the best of my knowledge, this study represents one of professional communication's first examinations of the impressions and consequences of written feedback in performance evaluations. Although the present study does have several methodological limitations, it suggests some important implications for 1) how to write appraisal feedback, 2) how to improve the training of raters, and 3) how to use rhetorical theory in creating evaluation instruments.

Discussion

The data strongly indicates the validity of the pilot study findings. There are specific words, phrases, and sentence structures that can make feedback more useful. Conversely, there are characteristics of certain kinds of feedback that are more than likely counterproductive to improving performance. This is important because research has established that feedback in evaluations is helpful to the evaluatee, yet little is currently understood about its mechanics. While there are no completely steadfast rules regarding feedback (in other words, it is unlikely that a phrase will be universally effective or ineffective), there are ways of writing feedback that tend to work better than others. We'll examine these by the three categories that were established during the pilot study, and then more closely examine the differences between evaluators/evaluatees and male/female respondents and the subsequent implications of those differences.

Word usage. One important discovery not highlighted in the analysis of hypotheses is that word usage may have a confusing effect on evaluation feedback. Survey items that contained a mixture of words appropriate and inappropriate to feedback received mixed ratings. Some respondents felt the feedback may improve performance, while others felt the feedback was likely to be counterproductive. It may be as important to avoid mixing word usage as it is to always frame feedback positively. But there was clear indication that phrases that are completely negative, "because you did not maintain authority I have never seen poorer classroom management" should be avoided. No respondents felt that such a phrase could improve performance. Likewise, a phrase devoid of negatives and offering the promise of improvement, "one way to improve classroom management would be to maintain authority," was universally selected to have the capability of improving performance. However, one reason that negative comments persist in performance evaluations is the evaluator's exigence in providing such feedback. Negative feedback may be viewed as an insurance policy against lawsuits in which a teacher has performed inappropriately and negative feedback may also be gathered as evidence for the decision to terminate an employee.

Content. Survey results indicated that the inclusion or exclusion of content in written feedback may affect performance. Respondents were more likely to mark feedback phrases in this section as neutral than in the other two categories. This is probably a result of the phrases being largely devoid of meaning. The one phrase (Table 6, Chapter 3) that did receive a majority "may improve performance" rating was in question form and indicated a specific part of teacher performance: accommodations

(plans prepared that help teachers responded appropriately to students with special behavioral and learning needs). However, even this phrase received a wide range of ratings. This is likely because it is unclear what about the classroom situation and teacher performance would necessitate reviewing accommodations. Several of the open response examples were much more thorough in their commentary and invoked a sense of collegiality between rater and ratee. It is probable that this sense of collegiality is associated with contextual understanding, and makes feedback (especially constructive criticism) easier to understand and follow.

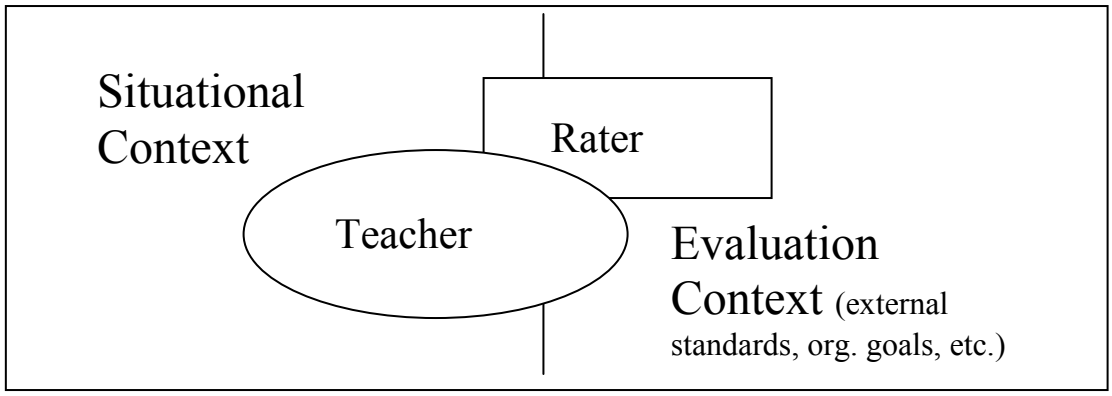
Structural order. Survey results in the structural order category established some tentative rules regarding forms of feedback phrases. The feedback framed in question form received a mixture of ratings, but the tendency was that this phrase would be counterproductive to improving performance. It is likely that the success/failure of a comment phrased as a question is contingent on the personality of the ratee. Such a phrase might feel condemning or accusatory to some individuals, even if the rater's intent was to simply make the rate consider alternatives. In other words, feedback in question form may appear judgmental, interfering with its usefulness. Additionally, results strongly indicated that the best practice for correcting performance problems is to offer the ratee a genuine compliment followed by a related suggestion for change. This structural order gives the appearance that the rater appreciates the ratee's work and is honestly trying to help them improve. Conversely, if such feedback is reversed, and the rater informs the ratee of what went wrong prior to giving a compliment, the feedback may be less effective.

Differences in Evaluators and Evaluatees. Some of the problems affecting evaluation feedback effectiveness could be a result of differences between evaluators and evaluatees. The two administrators in the sample often assessed feedback statements differently and sometimes counter to the majority of those that would be evaluated. For example, one administrator marked that a phrase containing "failed" in the word usage category may produce improvement in performance, while most of the teachers thought the phrase would adversely affect performance. This reflects a difference in approach to providing constructive criticism that can affect performance outcomes. It is likely that most administrators have good intentions for evaluation feedback, they may just not be aware of how to best provide it or constrained from providing such feedback. The study's findings should assist administrators in writing their evaluations.

Similarly, while part two of hypothesis two, that administrators and teachers do not share the same level of contextual understanding, was not confirmed, there is some evidence that neglecting to purposefully demonstrate awareness of the situation can lead to ineffective feedback. It is possible, then, given that administrators are not in the day-to-day classrooms that they are evaluating, that they do not always have an appropriate level of contextual understanding. It is much more likely that the rater's knowledge of context is largely grounded in the administrative processes of the organization, while the ratee's context is grounded in the observation setting which is, in this case, the classroom. This relationship is illustrated in Figure 3. The large box on the left represents the context of the situation in which the observation is taking place. The large box to the right of the situational context box represents the outside context of the

performance observation process. The rater falls mostly outside of the situational context because the rater has not spent enough time in the situation to develop a thorough understanding. The teacher has a much stronger understanding of the situational context. Conversely, the rater has a better understanding of the appraisal process than the teacher, which produces further disconnect between the teacher's and the rater's perception of performance observation. However, there is still some shared knowledge of context between the teacher and rater, as illustrated by the area in which the two overlap.

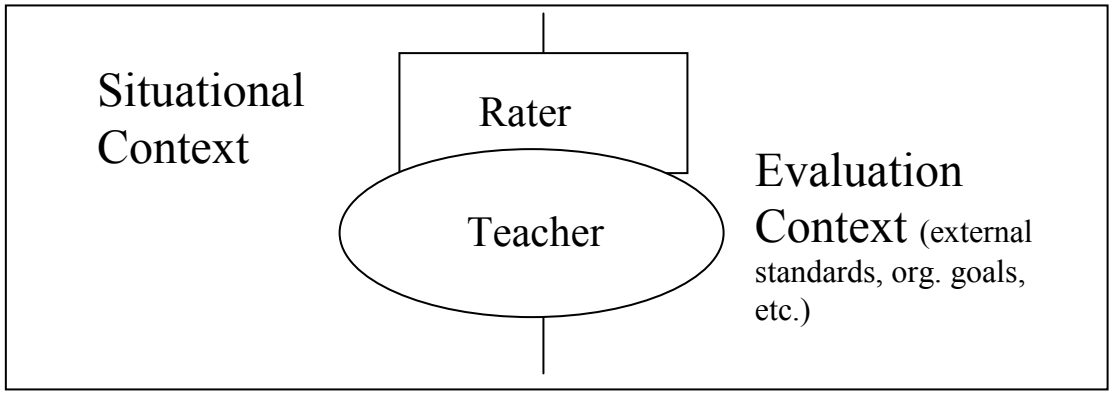
Figure 3
Actual Perception of Contexual Awareness between Ratee and Rater



However, the survey results demonstrate that it is possible to write feedback with a high level of contextual awareness even with a small amount of contextual knowledge. Several of the open responses to the short performance scenario in the content section of the survey exhibit contextual awareness. There is overt explanation of the situation and sympathetic reflection regarding the difficult task of keeping students productive on a Friday after a school event: "I can understand the need to maintain control on Fridays.

The end of the week is a tough time…" and "having been in your shoes…." These comments differ from the one administrator response, which makes no mention of the situation. These findings suggest that evaluation feedback is more effective if the rater's contextual awareness is perceived to mirror that of the teachers (Figure 4). Basic logic alone would leave one to conclude that the more an individual demonstrates understanding of one's plight in a given situation, the more receptive one would be to any discussion of that plight.

Figure 4
Ideal Perception of Contexual Awareness between Ratee and Rater



Future Directions for Research

The present study needs to be replicated and expanded on a much larger scale in to make more generalizable and valid conclusions regarding how language and the rhetorical situation impacts the effectiveness of written feedback. We especially need to obtain much more data on those who perform evaluations, as many of the expected relationships were unclear. It is possible that one of the causes of the low response rate to the survey was that most individuals within the targeted organizations are either

disinterested in or are against conducting performance evaluations. This opinion is not unwarranted, as this study demonstrates; evaluations can be quite unpleasant depending on how they are conducted. Replicating the study, then, would require much more discussion between potential organizations and the researcher. Perhaps the survey would receive a better reception after organizations have a better understanding of the intent of the study and the ratees were reminded that evaluations are likely to continue despite their often unpleasant nature. It would also be beneficial to follow up the survey with more qualitative assessments, such as think-aloud protocols where participants give their impressions of feedback as they read it, and interviews that would disclose more information about the appraisal process. Obtaining such data would greatly assist in constructing validity. But there are also other approaches to researching performance evaluation feedback that will broaden our present understanding of the subject matter. We will examine these by the three academic fields upon which this research draws.

Professional Communication. This research provides professional communicators with valuable insight regarding the rhetorical situation of a particular business genre-the performance evaluation. There is strong indication that there is a wide power differential between audience, subject, and writer. Those individuals being evaluated constitute an audience that holds little power within the rhetorical triangle despite holding a major stake in evaluation outcomes. They commonly have no control over the subject matter because the writer and the designer of the evaluation instrument have dictated that material. They generally have little influence on the evaluator aside from the overt

behaviors that provide the exigence for feedback, and neither the evaluator nor evaluatee may have had direct input in the design of the evaluation instrument.

This particular dilemma, a marginalized audience within the rhetorical situation, is not new to professional communication. In fact, it is because of professional communication's focus on the rhetorical situation and its components that make it ideally suited to be at the forefront of research in performance evaluations. For instance, Self and Self (1994) examine electronic communication practices that marginalize particular social and economic groups. The field also uncovers and explains differences in audiences and how that may affect communication practices (see Kostelnick, 1995 for a good example of how communication practices differ depending on the targeted audience). Additionally, professional communication scholars share much in common with the teaching performance that is the subject of this study: many of them are also educators and study communication practices within educational contexts. For example, Freire is well known in both fields for teaching others how to communicate.

Because professional communication has long studied written and oral communication, there are academic field specific tools that may assist in performance evaluation research. It would be beneficial to conduct textual analyses of actual performance evaluations. Through think-aloud protocols (research method in which participants are instructed to verbalize all thoughts as they process or write a text), one could learn more about the cognitive processes that occur as a rater writes an evaluation and how a ratee receives the feedback in an evaluation. Performing these kinds of

research would provide data that would help us to understand the reactions of the rater and ratee during the appraisal process.

Education. We need to better understand the mechanics of reacting to a performance evaluation, especially when the evaluation serves a developmental purpose. It is difficult to determine whether reactions to feedback are a function of the work being performed, of the evaluation purpose, or both. Educational scholars should further explore the unique experience of teaching as work performance. As Kyriakides, Demetriou, & Charalambous (2006) explain, to properly determine how to measure teaching performance, best theories and practices for teaching must first be agreed upon and implemented. This distances teaching "work" from other varieties of work considerably because of the diverse approaches teachers may take in order to educate students. As discussed in Chapter One, evaluating teaching is not as easy as evaluating the fast food worker for prompting a customer to order an additional item. While there are certainly some teaching practices that are almost universally agreed upon as ineffective or unacceptable (for instance, corporal punishment), it is much more likely that evaluators will encounter a variety of acceptable theories and practices in the instructional situation. Future studies should focus on identifying more classroom practices that are succeeding in the field so that administrators and evaluation instrument designers can incorporate them into the knowledge base with which they enter the rhetorical situation of evaluation.

Furthermore, in order to fully illustrate how feedback in evaluations of teaching may differ, the present study could be repeated for a different occupation or location. For

instance, workers in a more visceral field, such as assembly line workers, may place different values on their performance evaluation. This could lead to radically different perceptions of what constitutes appropriate feedback. Teachers working in areas that do not emphasize the evaluation as a guidepost for making decisions regarding future employment may not be as concerned with negative feedback. A comparative study that examines how teachers not in high stakes evaluation areas respond to feedback versus those teachers in areas where evaluation is integral to continued employment would help the discourse community in this study better gauge the significance of feedback.

Industrial/Organization Psychology. Finally, this study presents new approaches to semantic framing as understood by Industrial/Organizational Psychology. Traditionally, semantic framing has been grounded in a few conventional words and phrases (i.e. half empty/half full) that appear to have near universal appeal and effect. However, it is probable that the effect of framing is much more audience dependent than appears. Words and phrases that have multiple meanings or that are particular to the discourse community may convey different meaning in the evaluation context than in the common vernacular. Terministic screens particular to a specific discourse community could very well have unintended affects on the individual being evaluated. Future research should explore the difference between framing feedback in the community's vernacular or framing it in more common language.

Future Directions for Practice

This study leads me to make several recommendations that the educational community might implement in order to achieve best practices for performance evaluations.

First, there needs to be more recognition of the contributions that the professional communication field can make to the construction of evaluation theories and practices. This includes a solid grounding in rhetorical theories and communication research in order to situate professional communication as the leader in performance evaluation and implementation. Most importantly, professional communication offers user centered design techniques that can restore equilibrium to the rhetorical triangle and prevent the marginalization of the individuals being evaluated. The performance evaluation instrument is a functional document like other business genres. Thus it can benefit from usability testing, which can access how well the instrument serves its purpose. Feedback from all three stakeholders (school system authorities, the administrators that conduct evaluation, and the teachers being evaluated) in the evaluation process should be gathered in order to determine each party's needs (Duin, 1993). After gathering such information, it will be possible to create an instrument that promotes equality and effectiveness in feedback. If it is impractical to gather such data, document designers may purposefully use the scenario principle (Flower, et al., 1983) in order to envision the sort of instructional situations in which teachers and evaluators find themselves. Creating "paper people" or hypothetical users will allow designers to make accommodations for the variety of demographics that utilize the evaluation.

Second, there should be more emphasis on evaluator training in rhetorical skills. Evaluators must not only understand the evaluation instrument, but also understand how to respond to that instrument appropriately. Future training might include demonstrations of how to write effective constructive feedback and ways to avoid ineffective feedback. Evaluators must also be trained to identify contextual clues that might explain performance behaviors that are questionable or unique. Equipping evaluators with these tools should improve their side of the appraisal process.

There also should be more emphasis on training those being evaluated. The focus of this training is not instructional strategies or other ways to improve performance but rather to better inform the evaluatees of the process and its goals. There appears to be some ambiguity regarding the purpose of the evaluation, which may impact performance in unintended ways. Evaluatees might better understand how to interpret the evaluation and its feedback if they are educated about the evaluation process. For instance, if the evaluation was purely administrative (i.e., the purpose of the evaluation is recordkeeping, hiring, and/or termination), an evaluatee might not be concerned about the level of constructive criticism offered. Conversely, if the evaluation was developmental (i.e., the purpose of the evaluation is to improve classroom performance), a lack of constructive criticism would be of great concern.

In addition, school systems might consider implementing evaluator/evaluatee discussion groups. The intent of these discussion groups would be to discuss the evaluation process and inform both evaluators and evaluatees about the characteristics of their respective contexts. The groups would primarily discuss the responsibilities and

challenges of each side's task. For the teacher, this means informing the evaluator about specific situations and students in the classroom that dictate the teacher use particular educational strategies. For the evaluator, this means informing the teacher about the implications and intent of the appraisal, as well as the foci of the evaluation. However, the purpose of this exchange of information is not to provide a "cheat sheet" for the evaluatee. Instead, it is to inform both parties of their expectations for the evaluation.

Another recommendation is that schools systems implement an evaluation program that is not only top-down (i.e., the administrator evaluates the subordinate/teacher) but also includes evaluations from peers and students (often known as 360 degree appraisals). By having teachers and students evaluate other teachers, some of the contextual knowledge gap is eliminated. Students in a particular class are immersed in the material and teaching style. Other teachers provide valuable contextual knowledge because they are experiencing similar performance scenarios. These individuals would need to be oriented with the evaluation process in order to obtain validity, and it may be that some students are not able to provide accurate and honest feedback. However, this should not deter from the possibility of having a fresh perspective that provides unique and valuable feedback. For instance, student evaluations of instructors are required in colleges and universities and provide invaluable feedback about teaching and curriculum, as well as learning environments and learning outcomes.

Finally, while there is no standardized ADEPT evaluation form (districts currently have the option of using their own form), it appears that the evaluation instruments most

commonly used are not designed to encourage appropriate feedback. It is possible that this problem is a result of evaluations having dual purposes (to develop instructional practices and to make administrative decisions). An instrument that tries to accomplish both purposes may collapse upon itself: there is no motivation to provide developmental information if an administrative decision (such as termination of employment) is imminent. This reduces the evaluation to a catalog of evidence instead of a tool for improvement. The most effective way of resolving this issue is to simply conduct separate evaluations for both purposes. However, this may be unfeasible given the time commitment it would take to evaluate each teacher twice. Thus, it may be more practical to explore how the instrument may be modified in its current state.

There are very few visual and textual clues in the state-issued evaluation that indicate the need to provide any feedback, much less effective feedback. For example, the ADEPT Formal Evaluation Summary (see Appendix E) requires the evaluator to document evidence for each performance dimension. However, the instrument prompts the evaluator to document only *what occurred*, not to pronounce judgment. The evaluator's interpretation of whether or not the evaluatee fulfilled the requirements for a performance dimension is then inferred from two checkboxes, one that reads "met" and one that reads "not met." The "not met" option also requires the evaluator to check off on a statement that indicates the reason that the performance standard was not reached. These reasons are signified as ways to improve; however, items like "establishing rules and procedures" and "conveying an understanding of the students" fall short of providing meaningful feedback that could lead to improved performance.

In other words, there are no specific instructions or behaviors suggested regarding how or what rules and procedures to establish. This ambiguity could easily lead to disappointment on the part of the evaluator and confusion for the evaluatee. The evaluatee might institute new procedures for how students are to ask the teacher a question, but the evaluator was actually indicating that the teacher needed to revise the policy for turning in late work. This indicates the need for more flexibility in the design of the instrument so that evaluators have ample space to provide substantive feedback. For example, the checkboxes for areas of improvement might read "try establishing rules and procedures that address…," and include room for an open-ended response by the evaluator.

The area for documentation of evidence also could be modified to be more conducive to effective feedback. In addition to prompting the evaluator to record instances of each facet of the performance dimension, it could also prompt the evaluator to note exemplary and insufficient performance instances by leaving a description of the behavior/instance as well as a rationale for the judgment. Neither suggestion for revision of the instrument would require a substantial time investment, nor do they undermine the established criteria for evaluation. However, redesigning the evaluation instrument so that it overtly prompts the evaluator to mention what was done well, what was not done well, and how performance may be improved, would greatly increase the likelihood of providing feedback that is meaningful.

Summary

Professional communication's tools, theories, and techniques allow for an examination of performance appraisals not obtainable from other fields. It is well suited for becoming the leader in conducting this research because of its understanding of the rhetorical situation and its long experience in working with business documents. It can establish the rhetorical parameters for the performance evaluation creation and implementation and for analyzing the data from evaluation feedback.

This study begins the field's exploration of this research area by examining the language and rhetorical situation of performance evaluation feedback of a small sample of teachers. However, professional communication can assist in the creation of performance evaluation systems through its emphasis on user centered design. It can examine the rhetorical situation of evaluations in progress through think-aloud protocol and research past evaluations through textual analyses, all with an acute eye for the functionality of the communication process. It is my hope that this research establishes the need for professional communicators to enter this field of study and demonstrates to the stakeholders of evaluations that our discourse community might make evaluations more worthwhile, thus improving performance and attitudes in the workplace.

APPENDICES

APPENDIX A

Pilot Study Findings

My preliminary research regarding effective communication in performance appraisal comments suggests that there are ways raters can improve the effectiveness of written feedback. Specifically, they can improve evaluations by making comments that reflect proper use of content, word usage, and structural order. These preliminary conclusions are based on analysis of in-depth questionnaires completed by three student teachers in March 2006. The student teachers were all participating in the secondary education program at a respected southeastern private university. They had just finished their term of student teaching, which includes several evaluations from a variety of raters. Each student teacher completed the questionnaire (see Appendix A) at their discretion and submitted the completed version electronically.

The questionnaire was designed to solicit responses concerning effective communication in multiple ways. Participants responded to both actual and ideal situations, and were asked to reflect upon their personal experiences with how communication in performance evaluations succeeded or failed. My analysis of the questionnaire data found nine specific aspects that impact evaluation comment effectiveness. These nine aspects were then grouped into three categories: 1) *Word usage*-tone, passive voice, absolutes, and preciseness; 2) *Content*-evaluation purpose, uninformed comments, informed comments, and qualifiers; and 3) *Structural order*-sentence arrangement and syntax. What follows is an in-depth discussion of each category, its definition, and accompanying illustrative examples.

Word Usage

The first area important to producing effective communication is word usage. The most often mentioned aspect of this category is tone. Two of the student teachers noted that words should be chosen so that the evaluation does not appear overtly negative or positive. It is possible that this is because an overtly negative or positive tone might reduce the effectiveness of feedback- a ratee who receives all positives or negatives tends to dismiss the comments because of a perceived lack of validity. The second aspect of word usage relevant to effective communication is the use of passive voice. Selective use of passive voice can help soften the blow of negative criticism, while still getting the rater's point across. One student teacher reflects that words appropriate to evaluation comments are:

> "…words that allow for a little wiggle room, that do not definitively state an individual did not do well, but rather gently advise an individual to try something different…."

Closely related to the positive use of passive voice is the negative use of words that indicate a sense of absoluteness. Words such as "never" and "failed to" were indicated as inappropriate for feedback. While none of the student teachers indicated the reason for this, it might be inferred that words that imply absolutes are contradictory to the purpose

of the evaluation: words that don't allow room for change do not help to develop better teachers.

Content

I found that the student teachers emphasized content as one of the primary factors of effective communication in evaluations. Practically, there must be some kind of communication to be effective. The purpose of the evaluation should determine the content. In this case, all evaluations were for developmental purposes. Thus the content included should help the teacher improve as an educator. As one student teacher explains:

> "I have always considered performance observations as a method of giving constructive feedback purposed to improve teaching methods in a classroom. When this is kept in mind by the observer and the observed, I believe they can be very beneficial."

Clearly, for the performance evaluations of student teachers to be effective, comments should be written with the developmental goal in mind.

The second and third aspects of content, uninformed and informed remarks, demonstrate ways in which the rater can illustrate awareness of the situation. When the rater responds negatively because they are not fully informed of the instructional context, this leads to a lack of teacher respect for the performance instrument and a high degree of

inability to accept performance feedback, as well as make changes on that feedback. The student teachers' dislike of uninformed remarks and affinity toward informed remarks indicate that situational awareness is valued, and this can lead to more effective communication. The questionnaire respondents were less receptive to feedback when they felt like the rater did not make an effort to understand the situation. One student teacher reflects that

> "One observation a professor commented on the classroom management as if I was not aware that this particular class had not gone well…. They also did not take the time to learn the background of the classroom. It was an unusual day for those students and the observer did not take that into consideration, because they never took the time to ask about the students before observing."

Conversely, comments that appeared informed were thought of as beneficial:

> "These [comments] stood out to me because they were criticisms that I saw in my own teaching to be true, and the observer's suggestions gave me a new perspective on how to handle the areas in which I know I am weaker."

This indicates that part of effective communication in evaluations includes content that reflects that the rater has an adequate understanding of the situational context.

Closely rated to this issue is the final aspect of content: qualifiers. The inclusion of qualifying statements was viewed as integral to providing constructive feedback. Comments like "you did this *because*…" were highly favored over remarks that appeared based only on opinion. This was especially true for negative comments, as one participant stated:

> "…negative comments are pointless unless an observer offers some advice."

This speaks to the importance of stating explicit reasoning for evaluation comments, because they help to reinforce the other aspects of content (informed remarks and understanding of purpose) that lead to effective communication.

Structural Order

The third and final category that can affect the effectiveness of communication in performance evaluations is structural order. Structural order includes sentence arrangement and syntax. One student teacher recommended that evaluation comments start with positive feedback and then move on to constructive criticism, rationalizing that ratees are more receptive to criticism after being made to feel good. This suggests that comments should be arranged so that positive statements occur before the negative. However, unlike the other areas, there was not complete consensus regarding the arrangement of sentences. The other two student teachers suggested that comments should occur in either the chronological order that behaviors are observed or that they

should strictly follow the format of the evaluation instrument. One possible explanation for this is that the first student teacher's area is in the humanities whereas the other two student teachers both teach mathematics. While it is impossible to come to a definitive conclusion on this issue, it might be that the different approaches to this aspect are the result of a general difference in thought process. The first student teacher was more concerned with ratee emotion whereas the second and third teachers focused on procedures. Nonetheless, I suggest that, for general purposes, prefacing negative statements with positive ones can improve the effectiveness of communication, even if they occur within the procedural constructs suggested by the second and third teachers.

The syntax of individual statements is the second aspect of structural order impacting the communication in evaluations. Two of the three student teachers suggested that conciseness is present in most performance evaluation remarks. However, while "getting to the point" is important for clarity, the other aspects discussed in the report indicate that conciseness should not be done at the expense of providing adequate content. Additionally, one respondent adamantly opposed phrasing negative feedback as questions (i.e. "why did you do this?"). This is presumably because it appears accusatory, and not constructive.

This pilot study identified three areas key to effective communication in the performance evaluations of student teachers. Clearly, content, word usage, and structural order can be adapted and modified to meet the needs of the ratee. However, this study focused on *student* teachers, not fulltime teachers, and was a small convenience sample. Additionally, it did not address the perspective of the rater. Because it is not feasible to

draw definitive conclusions from this study, it is imperative to further pursue my research question in a manner that can test and expand upon these preliminary findings.

APPENDIX B

Pilot Study Questionnaire

Observation Questionnaire

Please fill out this basic information which will help me better understand your position and provide me with a way to contact you if necessary.

***Name**:*

Gender:

Subject Taught:

Email address:

Phone number (not required):

Instructions:

This questionnaire will ask about performance observations, but be aware that I'm not going to ask you to comment specifically about a rater (person conducting observation) or a particular observation. I ask that you do not provide names.

Before you start, I'd like for you to take a few minutes to reflect on performance observations that you've experienced. You might even gather some of your past observations. **This will help you to answer my questions better.**

Part A.

Actual Self. Now I'm going to ask you some questions about how you actually perceive performance reviews. Please write in the space provided below each question. Please answer the questions fully and to the best of your ability.

1. What is your general perception of the purpose of performance observations, and how well do they accomplish that purpose?

2. Describe for me the general sense of tone and wording (grammar, sentence structure, etc.) in performance observations.

3. Is there a general order to the way comments are made in performance observations? Describe that way.

4. How do performance observation comments usually begin and end? Why do you think that is?

5. Have you ever encountered any performance observation remarks that have stood out in your mind? Why did they stand out?

6. Are negative observation remarks often explicitly qualified, or are they made without written basis? ***(It might be useful to provide an example at this point of a qualified and unqualified remark, such as: You handled that student inappropriately vs. You handled that student inappropriately because you neglected to address his special needs.)*** *And why do you think this is?*

Part B.

Ideal Self. Now I'm going to ask you some questions about how you'd perceive an ideal performance review. Please write in the space provided below each question. Please answer the questions fully and to the best of your ability.

1. Ideally, what do you think the purpose of performance observations should be-is that purpose different from the actual purpose?

2. Could you give me a list of words you feel that are good words to use in performance observations? Why are these words appropriate for performance observations?

3. Could you give me a list of words you feel that are bad words to use in performance observations? Why are these words inappropriate for performance observations?

4. Do you prefer for negative observation remarks to be explicitly qualified, as was discussed earlier in the questionnaire? Why?

5. How should performance observation comments generally begin and end? Why?

Part C.

Others. Now I'm going to ask you some questions regarding your perception of how other individuals interact with performance reviews? Please write in the space provided below each question. Please answer the questions fully and to the best of your ability.

1. Do you feel that raters (people performing observation) share your sentiments regarding the purpose of performance observations? Why or why not?

2. Do you think that teachers and raters are generally happy with the way performance observations are written? Why or why not?

THE END

APPENDIX C

Recruitment Letter

Dear,
Hello, my name is Michael Hedges. I'm a graduate student in the MA in Professional Communication program at Clemson University. One of my areas of interest is in the rhetoric of evaluation feedback. I've been in contact with Dr. Kathryn Meeks, who oversees educator quality for the state of South Carolina, and understand that ADEPT program evaluations are good sources of written feedback.

Currently, I am in the process of conducting research on what makes language in feedback effective or ineffective. The goal of this research is to find ways to provide constructive criticism that provides maximum benefit for the evaluator and evaluatee. Dr. Meeks indicated that I would need to request permission from each South Carolina school district's superintendent, thus you are receiving this email because you are the contact I have for that position.

I believe that South Carolina's teachers and administrators would be an excellent participant pool for my current study-a short (10 min), anonymous, web-based survey that allows participants to critique words and phrases one might see in evaluations. This research would only require one recruitment email and one follow-up reminder email to teachers and administrators.

I plan to report my analysis of data to my research committee AND to the entities that participate in the research, making this a mutually beneficial project.

If you are interested in allowing me to conduct such a study or directing me to the appropriate contacts, please contact me and I'll provide more thorough information as well as allow you to review research materials.

Participation, on your part, only requires sending an invitation to the survey to the email listserv for district high schools.

Again, I think this is an excellent opportunity to learn how to make evaluations even more effective within the current system, and I sincerely appreciate your time and interest.

Michael Hedges
Candidate, MA in Professional Communication
Daniel Hall 219
Clemson University
864 656-6410 (office)
864 908-6005 (home)

APPENDIX D

Survey Invitation

Request / Consent to Participate

To: ####### Schools Listserv
From: Michael Hedges, Clemson University graduate student
Subject: Evaluation Feedback survey

If you either receive or give formal written feedback as part of ADEPT performance evaluations, please take a few minutes to fill out an online survey about ways in which feedback may be improved:
http://www.surveymonkey.com/s.asp?u=92583220322

PLEASE SUBMIT YOUR SURVEY RESPONSE BY FRIDAY, March 30.

I hope the results of this research will help both teachers and administrators who perform and receive evaluations improve their strategies for writing and receiving evaluation feedback.

Your participation is very important to the success of this research.
If you do not receive or perform ADEPT evaluations, please forward this request to the person(s) with that responsibility.

If you agree to participate, you will be asked to answer 13 questions about written feedback in performance evaluations. The survey should take about 10 minutes to complete.

Your participation is entirely voluntary; you may withdraw from the study at any time and you may decline to answer any question.

You will not be asked to provide your name or any contact information. If you have any questions about this study, or if any problems arise, please contact professor Cynthia Haynes at 864-656-3040 texcyn@clemson.edu or me at 864-908-6005 hedges@clemson.edu. If you have any questions or concerns about your rights as a research participant, please contact the Clemson University Office of Research Compliance at 864-656-6460 lmoll@clemson.edu.

By completing and submitting the survey, you give your informed consent to participate in this study. **I will make the study results available in June 2007 and send a message to your district to be distributed so you can see my findings.**

Thank you very much for your help!

Michael Hedges
Master of Arts in Professional Communication Candidate
219 Daniel Hall
Clemson University
Clemson, SC 29634-0101

tel: 864-656-6410
cel: 864-908-6005
hedges@clemson.edu

APPENDIX E

ADEPT Information

Overview of the ADEPT System (provided by South Carolina Board of Education)

June 2006

South Carolina's system for Assisting, Developing, and Evaluating Professional Teaching (ADEPT) grew out of the knowledge that good teaching is fundamental to student achievement. Implemented statewide in 1998, the ADEPT system has become a vital part of the state's overall teacher quality initiative. In addition to achieving the minimum score or better on appropriate examinations on both content and general teaching area knowledge, as established by the State Board of Education, educators must successfully complete all ADEPT requirements in order to be eligible to advance to a professional teaching certificate.

Central to the ADEPT system is a set of expectations for what teaching professionals should know, be able to do, and assume responsibility for accomplishing on an ongoing basis. These expectations, known as the ADEPT Performance Standards, are the linchpins that connect all stages of a teacher's career. A teacher's proficiency in each of the standards is expected to occur developmentally and to increase continuously throughout the entirety of his or her teaching career.

The ADEPT Performance Standards (APSs) for classroom-based teachers were revalidated in 2005 by a statewide committee made up of teachers, district and school administrators, higher education faculty, and representatives from professional

organizations. During this revalidation process, the committee's goal was to ensure that the APSs were consistent with current research and best practice. Following an extensive review of numerous national, state, and local teaching performance standards, including those from the Interstate New Teacher Assessment and Support Consortium (INTASC) and the National Board for Professional Teaching Standards (NBPTS), the APSs were revised and placed online for statewide field review. Additional modifications were made, based on feedback from the field, and the finalized APSs for classroom-based teachers were approved by the State Board of Education in June 2006. Separate

APSs for library media specialists, school guidance counselors, and speech-language therapists, respectively, were originally approved by the State Board of Education in 2003 and are slated for a later review cycle.

ADEPT Formal Evaluation Summary

Teacher's name _____ Contract level _____

District _____ School _____

Grade(s)/subject(s) _____ Academic year _____

Evaluation cycle ☐ Preliminary ☐ Final

Performance Dimensions for Classroom-Based Teachers		Consensus Judgment	
		Met	**Not Met**
PD 1	Long-Range Planning	☐	☐
PD 2	Short-Range Planning of Instruction	☐	☐
PD 3	Short-Range Planning, Development, and Use of Assessments	☐	☐
PD 4	Establishing and Maintaining High Expectations for Learners	☐	☐
PD 5	Using Instructional Strategies To Facilitate Learning	☐	☐
PD 6	Providing Content For Learners	☐	☐
PD 7	Monitoring and Enhancing Learning	☐	☐
PD 8	Maintaining an Environment That Promotes Learning	☐	☐
PD 9	Managing the Classroom	☐	☐
PD 10	Fulfilling Professional Responsibilities Beyond the Classroom	☐	☐
Overall Judgment An overall judgment of *met* indicates that the teacher meets expectations for at least nine of the ten performance dimensions at the time of the final evaluation.		**Met** ☐	**Not Met** ☐

Evaluators' signatures: By signing below, I verify that the formal evaluation process was conducted in accordance with the district's approved ADEPT plan and that I participated in making and am in agreement with the above judgments.

Evaluator _____ Date _____

Evaluator _____ Date _____

Evaluator _____ Date _____
(optional)

Teacher's signature: By signing below, I verify that I have received the results of this formal evaluation. My signature does not necessarily imply that I agree with these results.

Teacher _____ Date _____

Performance Dimension 1
Long-Range Planning

Key elements of the long range plan *typically* include

a description of the students that
- suggests a thorough understanding of the students' ability and developmental levels, backgrounds, needs, and interests; and
- includes appropriate sources that were used to obtain this information.

learning and developmental goals that are
- consistent with all appropriate curriculum standards;
- appropriately challenging for students' ability and developmental levels; and
- reflective of students' backgrounds, needs, and interests.

instructional units that are
- consistent with the appropriate curriculum standards and the long-range goals;
- logically sequenced in a timeline that provides sufficient opportunity to cover the key concepts and themes of the subject matter; and
- designed to expose students to a variety of intellectual, social, and cultural perspectives.

a description of key materials and resources
- that are needed to accomplish the goals; and
- that require special plans (including the description) for ordering or requesting.

procedures for assessing, evaluating, and recording students' progress and achievement that include appropriate
- assessment strategies,
- evaluation criteria, and
- documentation (record keeping).

rules and procedures
- for managing student behavior that are age-appropriate, stated in positive terms, specific, and consistent with best practice and relevant guidelines and regulations; and
- for handling non-instructional routines in an efficient manner.

procedures for routinely communicating with parents that are
- appropriate and
- designed to keep parents or guardians informed of and actively involved in the learning and developmental process.

Evidence Documentation for PD 1

The description of the students

The learning and developmental goals

The instructional units

The description of the key materials and resources

The procedures for assessing, evaluating, and recording students' progress and achievement

The rules and procedures

The procedures for routinely communicating with parents

Consensus Judgment for PD 1

☐ **Met**	☐ **Not Met**
The teacher develops an overall plan for facilitating student achievement of appropriate curriculum standards and long-range learning and developmental goals that includes a general organizational framework for progressing through the school year in an effective and efficient manner.	The teacher does not demonstrate long-range planning skills that are adequate to promote student achievement of appropriate curriculum standards and long-range learning and developmental goals. Suggested areas for improvement include *(check all appropriate)* ☐ conveying an understanding of the students ☐ establishing learning and developmental goals ☐ developing instructional units ☐ determining key materials and resources ☐ determining assessment and record keeping methods ☐ establishing rules and procedures ☐ determining essential communications ☐ other:

Performance Dimension 2
Short-Range Planning of Instruction

Key elements of short-range instructional planning *typically* include

an instructional unit (interview questions 1 and 2) that
- is integrated with other units and/or subject areas listed in the LRP, and
- reflects the teacher's understanding of the students' ability and developmental levels, learning styles, needs, social/cultural backgrounds, and interests.

learning and/or developmental objectives (interview question 3) that are
- consistent with the appropriate curriculum standards and with the long-range learning and developmental goals, and
- appropriately challenging for all students.

content (interview question 4) that
- is current and accurate,
- facilitates the students' achievement of the curriculum standards and the unit objectives, and
- exposes students to a variety of intellectual, social, and cultural perspectives as well as "real-life" situations

instructional strategies (interview question 5) that
- are appropriately matched to the objectives and to the content;

- are designed to actively engage students in instruction and learning;
- are designed to promote varied levels of thinking and problem-solving;
- are varied and logically sequenced to provide sufficient opportunities for initial learning, application, practice, and review;
- are designed to accommodate differences in rates of learning and development; and
- include appropriate materials, resources, and technologies.

a reflection on planning (interview question 9) indicating that the teacher

- is able to draw appropriate conclusions regarding any revisions that may need to be made to enhance instructional effectiveness.

Evidence Documentation for PD 2

The short-range plans for the instructional unit

The short-range learning and/or developmental objectives

The content

The instructional strategies

The teacher's reflection on planning

Consensus Judgment for PD 2

☐ **Met**	☐ **Not Met**
The teacher demonstrates effective short-range planning skills that enable him or her to design a series of integrated lessons or days of instruction in a manner that facilitates students' accomplishment of appropriate curriculum standards and learning objectives.	The teacher's short-range instructional planning skills do not adequately support students' accomplishment of appropriate curriculum standards and learning objectives. Suggested areas for improvement include *(check all appropriate)* ☐ developing the instructional unit ☐ establishing learning and/or developmental objectives ☐ determining content ☐ selecting instructional strategies ☐ reflecting on instruction to guide future planning ☐ other:

Performance Dimension 3
Short-Range Planning, Development, and Use of Assessments

Key elements of assessment *typically* include

a variety of formal and informal assessment methods (interview question 6) that are

- appropriate indicators of students' knowledge and skills associated with each objective;
- consistent with the content and the instructional strategies;
- appropriately selected, adapted, and/or developed;
- appropriately administered and/or applied;
- administered and/or applied as frequently as necessary to appropriately monitor both student progress and student achievement relative to each objective.

evaluation criteria (interview question 7) that

- allow for valid and reliable interpretations of student performance, and
- are appropriately high for students' ability and developmental levels.

use of assessment results (interview question 8) indicating that the teacher

- accurately analyzes and interprets student performance,
- provides appropriate instructional feedback to the students,
- provides appropriate assistance to the students, as necessary,
- maintains accurate and well-organized records of students' progress and achievement, and
- determines the extent to which the instruction is meeting the needs of all students.

Evidence Documentation for PD 3

The formal and informal assessment methods

The evaluation criteria

The teacher's use of the assessment results

Consensus Judgment for PD 3

☐ **Met**	☐ **Not Met**
The teacher appropriately gathers and analyzes student performance data, makes appropriate determinations regarding student progress and achievement, and uses this information effectively to guide subsequent instruction.	The teacher's short-range planning, development, and/or use of assessments does not adequately support instruction. Suggested areas for improvement include *(check all appropriate)* ☐ selecting and/or applying assessment methods ☐ establishing evaluation criteria ☐ analyzing, interpreting, and/or using assessment results ☐ other:

Performance Dimension 4
Establishing and Maintaining High Expectations for Learners

Key elements of establishing and maintaining high expectations for learners *typically* include the teacher's ability and disposition to

establish learning and developmental objectives that

- reflect appropriate curriculum standards,
- are appropriately challenging for students' ability and developmental levels,
- are clearly communicated to the students in terms of relevance and importance, and
- are clearly linked to students' previous and future learning.

establish and clearly communicate expectations to students regarding

- class participation and
- completing instructional activities, events, and assignments.

foster a sense of purpose and student responsibility for learning.

Evidence Documentation for PD 4

The teacher's expectations for student learning

The teacher's expectations for student participation

The teacher's expectations for student responsibility for learning

Consensus Judgment for PD 4

☐ **Met**	☐ **Not Met**
The teacher establishes, clearly communicates, and maintains appropriate expectations for student learning, participation, and responsibility.	The teacher does not adequately establish, communicate, and/or maintain appropriate expectations for students. Suggested areas for improvement include *(check all appropriate)* ☐ establishing appropriate learning and/or developmental objectives ☐ conveying appropriate student expectations ☐ fostering a sense of purpose and student responsibility ☐ other:

Performance Dimension 5
Using Instructional Strategies to Facilitate Learning

Key elements of using instructional strategies to facilitate learning *typically* include the teacher's ability and disposition to

- **use a variety of appropriate strategies effectively.**
- **promote varied levels of thinking and problem-solving.**
- **provide opportunities for initial learning, application, practice, and review.**
- **accommodate differences in students' ability/developmental levels, rates of learning, and styles of learning.**
- **actively engage students in instruction and learning.**
- **promote both independent and collaborative learning.**
- **promote positive and productive interactions between the teacher and the students.**

Evidence Documentation for PD 5

The variety of instructional strategies

The appropriateness of the instructional strategies

The effectiveness of the instructional strategies

Consensus Judgment for PD 5

☐ **Met**	☐ **Not Met**
The teacher promotes learning through the effective use of appropriate instructional strategies.	The teacher's selection and/or use of instructional strategies does not adequately facilitate student learning. Suggested areas for improvement include *(check all appropriate)* ☐ using appropriate strategies effectively

		☐ promoting varied levels of thinking and problem-solving ☐ providing opportunities for all stages of learning ☐ accommodating student needs and differences ☐ actively engaging students in instruction and learning ☐ promoting independent and collaborative learning ☐ promoting positive and productive classroom interactions ☐ other:

Performance Dimension 6
Providing Content for Learners

Key elements of providing content for learners *typically* include the teacher's ability and disposition to

select content that

- is aligned with appropriate curriculum standards,
- is appropriate for the learning and/or developmental objectives,
- is current and accurate,
- includes all important concepts and/or skills (and, conversely, eliminates unnecessary or superfluous information), and
- is derived from a variety of appropriate sources.

present content

- by identifying key concepts,
- using a variety of clear, logical, and appropriate explanations, examples, demonstrations, etc.,
- by relating the content to the students' present and future interests and needs,
- by stimulating critical and reflective thinking,
- at an appropriate pace.

Evidence Documentation for PD 6

The teacher selected content that

The teacher presented the content

Consensus Judgment for PD 6

☐ **Met**		☐ **Not Met**
The teacher demonstrates a thorough command of the subject matter relative to the appropriate curriculum standards and student learning objectives.		The teacher does not provide appropriate and/or adequate content to the students. Suggested areas for improvement include *(check all appropriate)* ☐ selecting appropriate content ☐ presenting content effectively ☐ other:

Performance Dimension 7
Monitoring and Enhancing Learning

Key elements of monitoring and enhancing learning *typically* include the teacher's ability and disposition to

use appropriate informal assessment strategies (e.g., observing, questioning, listening, reviewing student work)

- in an effective manner.

- on a continuous basis.
- to tap various levels of thinking and performance.
- with all students.

use appropriate formal assessment strategies (e.g., quizzes, homework, individual/group presentations, tests).
provide appropriate and prompt oral and written instructional feedback to students.
promote students' ability and willingness to monitor and evaluate their own progress.
adjust instruction to review, re-teach, or extend key concepts, as appropriate.

Evidence Documentation for PD 7

The informal assessment strategies

The formal assessment strategies

The teacher's feedback to students

The teacher's ability to involve students in monitoring and evaluating their own progress

The teacher's ability to adjust instruction, as needed,

Consensus Judgment for PD 7

☐ **Met**	☐ **Not Met**
The teacher uses dynamic assessment to facilitate learning for all students.	The teacher does not effectively use appropriate techniques to monitor and enhance student learning. Suggested areas for improvement include *(check all appropriate)* ☐ using appropriate informal assessment strategies ☐ using appropriate formal assessment strategies ☐ providing appropriate and prompt feedback to students ☐ helping students monitor and evaluate their own progress ☐ adjusting instruction, as needed ☐ other:

Performance Dimension 8
Maintaining an Environment That Promotes Learning

Key elements of maintaining an environment that promotes learning *typically* include the teacher's ability and disposition to

create a safe, stimulating, and inviting learning environment.
demonstrate confidence and enthusiasm.
maintain positive, respectful, and appropriate interactions with the students.
encourage students' active engagement in learning.
promote positive, learning-focused interactions, collaboration, and teamwork among the students.
provide appropriate extrinsic and intrinsic incentives for learning and rewards for progress and success.

Evidence Documentation for PD 8
The learning environment The teacher's level of confidence and enthusiasm The classroom climate The interactions The learning incentives

Consensus Judgment for PD 8	
☐ **Met** The teacher creates and maintains a classroom environment and climate that encourage and support student learning.	☐ **Not Met** The overall classroom environment and/or climate is not conducive to student learning. Suggested areas for improvement include *(check all appropriate)* ☐ creating a safe, stimulating, and inviting learning environment ☐ demonstrating confidence and enthusiasm ☐ maintaining appropriate interactions with students ☐ encouraging students' active engagement in learning ☐ promoting appropriate interactions among students ☐ providing appropriate incentives and rewards ☐ other:

Performance Dimension 9
Managing the Classroom

Key elements of managing the classroom *typically* include the teacher's ability and disposition to

- **establish efficient routines for completing essential non-instructional tasks.**
- **effectively manage instructional time, materials, resources, technologies, and transitions.**
- **establish, communicate, and enforce rules and procedures that maximize the occurrence of appropriate student behaviors.**
- **effectively manage any inappropriate student behaviors.**
- **promote students' ability and willingness to assume responsibility for their own behaviors.**

Evidence Documentation for PD 9
The routines for completing non-instructional tasks The teacher's management of instructional time, materials, resources, technologies, and transitions

The teacher's management of student behavior

The teacher's ability to promote students' ability and willingness to assume responsibility for their own behaviors

Consensus Judgment for PD 9

☐ **Met**	☐ **Not Met**
The teacher maximizes instruction by effectively managing student behavior, instructional routines and materials, and essential non-instructional tasks.	The teacher's management of student behavior, instructional routines and materials, and essential non-instructional tasks significantly compromise the teaching and learning processes. Suggested areas for improvement include *(check all appropriate)* ☐ establishing routines for completing essential non-instructional tasks ☐ managing instructional routines and materials ☐ maximizing appropriate student behaviors ☐ managing inappropriate student behaviors ☐ helping students assume responsibility for their own behaviors ☐ other:

Performance Dimension 10
Fulfilling Professional Responsibilities Beyond the Classroom

Key elements of fulfilling professional responsibilities *typically* include the teacher's ability and disposition to

establish, maintain, and be a contributing partner in collaborative professional relationships with parents, other teachers, other student-oriented professionals in the school and/or district, and members of the community in order to

- plan and provide appropriate learning experiences for students,
- accomplish school goals, and
- support the district's strategic plan.

meet all professional expectations in an ethical and responsible manner.

engage in continuous professional growth and development.

Evidence Documentation for PD 10

In terms of being a contributing partner in collaborative professional relationships, the teacher

In terms of meeting all professional expectations in an ethical and responsible manner, the teacher

In terms of engaging in continuous professional growth and development, the teacher

Consensus Judgment for PD 10	
☐ **Met** The teacher is an ethical, responsible, contributing, self-motivated, and ever-learning member of the profession.	☐ **Not Met** The teacher does not consistently demonstrate behaviors that are commensurate with professional expectations. Suggested areas for improvement include *(check all appropriate)* ☐ establishing and maintaining collaborative professional relationships ☐ meeting all professional expectations in an ethical manner ☐ meeting all professional expectations in a responsible manner ☐ engaging in continuous professional growth and development ☐ other:

APPENDIX F

Online Survey

Evaluation Feedback Survey Quit Survey>>

1. Welcome

This study is approved by the Clemson University Institutional Review Board.

*** 1. The following survey is designed to identify how the language used in performance evaluations may affect the likelihood of *producing change* in performance. If you agree to participate, you will be asked to answer 13 questions about written feedback in performance evaluations. The survey should take about 10 minutes to complete. Your participation is entirely voluntary; participants may withdraw from the study at any time without penalty and you do not have to answer every survey item. There is no penalty for not participating. Your participation is very important to the success of this research. You will not be asked to provide your name or any contact information. School's and individuals identities will remain strictly anonymous and confidential. If you have any questions about this study, or if any problems arise, please contact Dr. Cynthia Haynes at 864-656-3040 texcyn@clemson.edu or me at 864-908-6005 hedges@clemson.edu. If you have any questions or concerns about your rights as a research participant, please contact the Clemson University Office of Research Compliance at 864-656-6460 lmoll@clemson.edu. If you do not receive or perform ADEPT evaluations, please forward this request to the person(s) with that responsibility. By selecting "I agree" you imply informed consent to participate in the survey. Then, click "continue" to proceed.**

I agree

Next >>

Evaluation Feedback Survey Quit Survey>>

2. Word Usage

Paying close attention to *word choice*, rate the statements below by selecting the best rating from the drop-down menu below each item.

2. "You failed to manage the classroom properly because you did not maintain authority."

3. "You might improve your classroom management by maintaining authority."

4. "Because you did not maintain authority, I have never seen poorer classroom management."

5. "One way to improve classroom management would be to maintain authority."

6. In the space provided below, please list any specific words or phrases often found in performance evaluation feedback that you consider *counterproductive* to improving performance:

<< Go Back Next >>

Evaluation Feedback Survey

Quit Survey>>

3. Content

Read the following performance evaluation scenario:

Mrs. Lopez is in her fifteenth year of teaching Government. She tells you that her class has several children that need special accommodations and have trouble focusing without direct instruction. On the day that you are observing, the class is very talkative because they just came back from a school pep rally. Mrs. Lopez decides to assign a review worksheet that will be due at the end of class. The talking ceases. It is the last class on a Friday.

Paying close attention to the *content* of each feedback comment, rate the statements below by selecting the best rating from the drop-down menu below each item.

7. "Try other methods to achieve classroom management."

8. "While you were doing your lesson plan, what were your accommodations?"

9. "Good way to quiet down your kids!"

10. "Nice way to offer a review of material."

11. If you were writing the evaluation for Mrs. Lopez, what kind of feedback would you give her regarding this situation. Please type your response:

<< Go Back Next >>

Evaluation Feedback Survey Quit Survey>>

4. Structural Order

Paying close attention to the *grammatical structure and order* of each sentence, rate the statements below by selecting the best rating from the drop-down menu below each item.

12. "Why did you discipline student A that way?"

13. "You handled your class well. Perhaps, next time, you will consider disciplining student A by...."

14. "You did not discipline student A appropriately. But you handled the class well."

<< Go Back Next >>

Evaluation Feedback Survey Quit Survey>>

5. Other Info

Congratulations! You're almost finished.

Please answer the following questions by using the drop-down menu provided and then you'll be done.

*** 15. Which best describes your profession?**

16. If you are a teacher, which best describes your subject area?

17. Which best describes how long you have been employed in your current position?

18. Which best describes your gender?

<< Go Back Next >>

Evaluation Feedback Survey Quit Survey>>

6. You're done!

That's all, thanks for your help!

I will make the study results available in May 2007 and send a message to your district to be distributed so you can see my findings.

<< Go Back Finish >>

REFERENCES

Beran, T., Violato, C., Kline, D., & Frideres, J. (2005). The utility of student ratings of instruction for students, faculty, and administrators: a "consequential validity" study. *Canadian Journal of Higher Education*, *35*(2), 49-70.

Bergland, B. (1997). Incorporating performance reviews into the business writing classroom. *Business Communication Quarterly*, *60*(4), 52-62.

Bitzer, L. F. (1968). The rhetorical situation. *Philosophy and Rhetoric*, *1*(1), 1-14.

Burke, K. (2001). From *language as symbolic action* (1966). In P. Bizzell & B. Herzberg (Eds.), *The rhetorical tradition: readings from classical times to the present* (pp. 1340-1347). Boston: Bedford/St. Martin's.

Carliner, S. (2003). Characteristic-based, task-based, and results-based: three value systems for assessing professionally produced technical communication products. *Technical Communication Quarterly*, *12*, 83-100.

Curtis, A. B., Harvey, R. D., & Ravden, D. (2005). Sources of political distortions in performance appraisals. Appraisal purpose and rater accountability. *Group & Organization Management*, *30*, 42-60.

Davis, W. D., Carson, C. M., Ammeter, A. P., & Treadway, D. C. (2005). The interactive effects of goal orientation and feedback specificity on task performance. *Human Performance*, *18*, 409-426.

Dulewicz, V. (1989). *Performance appraisal and counselling*, in Herriot, P., *Assessment and selection in organizations: methods and practices for recruitment and appraisal* (pp. 645-649). New York: John Wiley & Sons.

Duin, A. H. (1993). Test drive-techniques for evaluating the usability of documents. In C. M. Barnum & S. Carliner (Eds.), *Techniques for technical communicators* (pp. 306-335). New York: Macmillan.

Dymoke, S., & Harrison, J. (2006). Professional development and the beginning teacher: issues of teacher autonomy and institutional conformity in the performance review process. *Journal of Education for Teaching*, *32*(1), 71-92.

Flower, L., Hayes, J. R., & Swartz, H. (1983). Revising functional documents: the scenario principle. In P. V. Anderson, R. J. Brockmann, & C. R. Miller (Eds.), New essays in technical and scientific communication: research, theory, practice (pp. 41-58). Farmingdale, NY: Baywood.

Fuller, J. A., Stanton, J. M., Fisher, G. G., Spitzmuller, C., Russell, S. S., & Smith, P. C. (2003). A lengthy look at the daily grind: Time series analysis of events, mood, stress, and satisfaction. *Journal of Applied Psychology*, *88*, 1019-1033.

Goldstein, J., & Noguera, P. A. (2006). A thoughtful approach to teacher evaluation. *Educational Leadership*, *63*(6), 31-37.

Goodall Jr., H. L., Wilson, G. L., & Waagen, C. L. (1986). The performance appraisal interview: an interpretive reassessment. *Quarterly Journal of Speech*, *72*, 74-87.

Gorgias. (2001). Encomium of Helen (ca. 414 B.C.E.). In P. Bizzell & B. Herzberg (Eds.), *The rhetorical tradition: readings from classical times to the present* (pp. 44-46). Boston: Bedford/St. Martin's.

Gueldenzoph, L. E., & May, G. L. (2002). Collaborative peer evaluation: best practices for group member assessments. *Business Communication Quarterly*, *65*, 9-20.

Holland, N. (1979). Reading and identity: a psychoanalytic revolution. *Academy Forum, 23*(2), 7-9.

---------,---. (1975). Unity identity text self. *PMLA*, *90*, 813-822.

Introduction to performance appraisal. (n.d.). Retrieved June 23, 2007, from Archer North's performance-appraisal.com: the complete online guide: http://www.performance-appraisal.com/intro.htm

Jarrett, C. (2005). Caroline's corner: one facilitator good, four facilitators better? Retrieved June 23, 2007, from UsabilityNews.com: http://www.usabilitynews.com/news/article2268.asp

Jawahar, I. M. (2005). Do raters consider the influence of situational factors on observed performance when evaluating performance? Evidence from three experiments. *Group & Organization Management*, *30*, 6-41.

Kain, D. J. (2005). Constructing genre: a threefold typology. *Technical Communication Quarterly*, *14*, 375-409.

Kyriakides, L., Demetriou, D., & Charalambous, C. (2006). Generating criteria for evaluating teachers through teacher effectiveness research. *Educational Research*, *48*(1), 1-20.

Latour, B., & Woolgar, S. (1986). *Laboratory life: the construction of scientific facts*. Princeton, NJ: Princeton University Press.

MacNealy, M. S. (1999). *Strategies for empirical research in writing*. Boston: Allyn and Bacon.

Manshor, A. T., & Kamalanabhan, T. J. (2000). An examination of raters' and ratees' preferences in process and feedback in performance appraisal. *Psychological Reports*, *86*, 203-214.

Marshall, K. (2005). It's time to rethink teacher supervision and evaluation. *Phi Delta Kappan*, *86*, 727-735.

Moore, S., & Kuol, N. (2005). Students evaluating teachers: exploring the importance of faculty reaction to feedback on teaching. *Teaching in Higher Education*, *10*(1), 57-73.

Murphy, K. R., & Cleveland, J. N. (1995). *Understanding performance appraisal: Social,organizational, and goal-based perspectives*. Thousand Oaks, CA: Sage.

Murphy, K. R., Cleveland, J. N., Skattebo, A. L., & Kinney, T. B. (2004). Raters who pursue different goals give different ratings. *Journal of Applied Psychology*, *89*, 158-164.

Myers, G. (1985). Texts as knowledge claims: the social construction of two biology articles. *Social Studies in Science*, *15*, 593-630.

Ornatowski, C. M. (1995). The writing consultant and the corporate/industry culture. *Journal of Business and Technical Communication*, *9*, 446-461.

Rogelberg, S. G., Church, A. H., Waclawski, J., & Stanton, J. M. (2002). Organizational survey research. In S. G. Rogelberg (Ed.) *Handbook of research methods in industrial and organizational psychology* (pp. 141-160). Malden, MA: Blackwell.

Rose, L. (2004). No Child Left Behind: The mathematics of guaranteed failure. NCLB: failed schools--or failed law?. *Educational Horizons*, *82*(2), 121.

Rose, M. (1996). *The promise of public education in America: possible lives*. New York: Penguin Books.

Rude, C. D. (2004). The report for decision making: genre and inquiry. In J. Johnson-Eilola & S. Selber (Eds.), *Central works in technical communication* (pp. 70-90). New York: Oxford University Press.

Scullen, S. C., Mount, M. K., & Judge, T. A. (203). Evidence of the construct validity of developmental ratings of managerial performance. *Journal of Applied Psychology*, *88*, 50-66.

Selfe, C. L. & Selfe, R. J. (1994). The politics of the interface: power and its exercise in electronic contact zones. *College Composition and Communication*, *45*, 480-504.

Smither, J. W., London, M., & Riley, R. R. (2005). Does performance improve following multisource feedback? A theoretical model, meta-analysis, and review of empirical findings. *Personnel Psychology*, *58*, 33-66.

Sturman, M. C., Cheramie, R. A., & Cashen, L. H. (2005). The impact of job complexity and performance measurement on the temporal consistency, stability, and test-retest reliability of employee job performance ratings. *Journal of Applied Psychology*, *90*, 269-283.

Tucker, P. D., Beers C. S., Strong J. H., & Gareis, C. R. (2003). The efficacy of portfolios for teacher evaluation and professional development: do they make a difference? *Educational Administration Quarterly*, *39*, 572-602.