

Old Dominion University

ODU Digital Commons

Educational Foundations & Leadership Theses
& Dissertations

Educational Foundations & Leadership

Summer 2010

Perceptions of Virginia Community College System Faculty and Administrators on the Purposes for and Composition of a Comprehensive Evaluation System for Teaching Faculty Members

William H. Hightower Jr.
Old Dominion University

Follow this and additional works at: https://digitalcommons.odu.edu/efl_etds



Part of the [Community College Education Administration Commons](#), [Educational Assessment, Evaluation, and Research Commons](#), [Higher Education Commons](#), and the [Higher Education and Teaching Commons](#)

Recommended Citation

Hightower, William H.. "Perceptions of Virginia Community College System Faculty and Administrators on the Purposes for and Composition of a Comprehensive Evaluation System for Teaching Faculty Members" (2010). Doctor of Philosophy (PhD), Dissertation, Educational Foundations & Leadership, Old Dominion University, DOI: 10.25777/gmzg-pg54
https://digitalcommons.odu.edu/efl_etds/126

This Dissertation is brought to you for free and open access by the Educational Foundations & Leadership at ODU Digital Commons. It has been accepted for inclusion in Educational Foundations & Leadership Theses & Dissertations by an authorized administrator of ODU Digital Commons. For more information, please contact digitalcommons@odu.edu.

PERCEPTIONS OF VIRGINIA COMMUNITY COLLEGE SYSTEM FACULTY AND
ADMINISTRATORS ON THE PURPOSES FOR AND COMPOSITION OF A
COMPREHENSIVE EVALUATION SYSTEM FOR TEACHING FACULTY

MEMBERS

by

William H. Hightower, Jr.
B.A. June 1979, University of California at Los Angeles
M.S. June 1981, The Ohio State University

A Dissertation Submitted to the Faculty of
Old Dominion University in Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY

COMMUNITY COLLEGE LEADERSHIP

OLD DOMINION UNIVERSITY

June 2010

Approved by;

Linda Bol, Ph.D. (Chair)

Molly H. Duggan, Ph.D. (Member)

Monty E. Sullivan, Ed.D. (Member)

ABSTRACT

PERCEPTIONS OF VIRGINIA COMMUNITY COLLEGE SYSTEM FACULTY AND ADMINISTRATORS ON THE PURPOSES FOR AND COMPOSITION OF A COMPREHENSIVE EVALUATION SYSTEM FOR TEACHING FACULTY MEMBERS

William H. Hightower, Jr.
Old Dominion University, 2010
Director: Dr. Linda Bol

A survey instrument was developed to measure community college faculty and administrator views on the faculty evaluation process. Responses were then compared based on demographic characteristics such as primary area of instruction, supervisory responsibility, years of experience, and gender. Open-ended survey questions asked respondents to identify the strengths, limitations and changes needed for their current faculty evaluation plans.

A total of 404 faculty members and 67 administrators completed the survey. Significant differences were found between faculty and administrator responses with respect to the reasons for conducting faculty evaluations, the relative importance of including certain elements in the evaluation process, and what data sources should be used in the construction of the final evaluation rating. Administrators consistently rated items related to faculty evaluation uses, inclusion of various evaluative elements, and the use of several data sources significantly higher than did the faculty members. Demographic differences were also discovered between faculty members when their group responses were compared based on primary area of instruction (transfer versus

career and technical education), years of full-time teaching experience, and gender. Transfer faculty rated external evaluation and service to the college significantly higher than their career and technical (CTE) colleagues, but rated the use of alternative instructional delivery formats and student performance significantly lower than CTE faculty. Faculty with less than 7 years of experience rated administrator and external evaluation significantly higher, and preparation for class significantly lower, than their more experienced colleagues. Female faculty members rated the use of alternative instructional delivery formats significantly higher than did male faculty members.

Qualitative findings supported the use of multiple measures of faculty performance, including student evaluation, supervisor evaluation, and reflective faculty narrative and/or portfolio self-evaluation. Respondents cited the need for objective, standardized criteria for evaluating faculty member performance, with teaching allotted the greatest weight. They suggested faculty evaluation should be primarily a formative process tied to professional development, and merit pay should be uncoupled from the formative evaluation rating and should be the result of a separate, competitive process. These results have implications for revising the current faculty evaluation process used in the Virginia Community College System.

DEDICATION

My most heart-felt thanks go to one fellow traveler in particular: Helen, my wife, my cohort-mate, the love of my life, and my biggest supporter. We met on the first day of the first summer institute and the rest has been a long, strange trip indeed. This is for you, Helen. When you cross the finish line I'll be right there with you...then we'll take off on a long overdue vacation!

ACKNOWLEDGEMENTS

“A journey of a thousand miles begins with a single step.” - Lao-tzu, 6th Century BC

“What a long, strange trip it’s been.” – The Grateful Dead, 1970

As with many journeys, the pathway between beginning a doctoral program and the completion of this dissertation often seemed like a journey of at least a thousand miles, full of twists and turns, roadblocks, delays, u-turns, and hazards. To say that it has been a long, strange trip is an understatement.

When a group of diverse individuals from across the Commonwealth of Virginia embarked as a cohort of fellow travelers on the maiden journey of the ODU Community College Leadership Ph.D. program in 2003, who could have predicted the twists and turns the group members would experience along the way? We shared many common challenges and triumphs along the way, particularly in surviving the first summer institute (or boot camp in Powhattan Village) and continuing in the program as “survivors” who helped each other through the coursework. That first summer set the stage for the rest of the journey and served as a bonding experience that joined all of us together for the rest of our lives – even as our individual paths weave, wind, and occasionally cross. Some members of the cohort treated the journey as a sprint (Ruth, Kellie, Pat, Martha, & Janet), while others took a more leisurely approach (Jay). Some of us took many by-ways and detours along the way (Dick, Joe, and me) but we never lost sight of the finish line of our marathon journey. Others are still somewhere on the road.

One step at a time, whether fast or slow, we reached the same destination - the completion of the dissertation. But we did not get here all on our own and I would like to

thank the “road crew” who helped me find the trail when I got off of the beaten path, stayed too long on a detour, or simply procrastinated at a rest stop along the way.

First, I would like to thank my dissertation chair, Linda Bol, who was one of the first faculty members I met during that fateful first summer institute. When she gave us two weeks to research and write a 20-page paper, I knew that I had found my dissertation chair – a personable taskmaster who enjoyed statistics and research methods! I had no doubt that she would make sure that I conducted a study that made a meaningful contribution to the literature. Anything done well in this dissertation is due to her influence. If there are any errors it is because I slipped in something when she wasn’t looking. She was understanding when I took extended periods away from working on the dissertation, but pushed me with aggressive timelines when I was ready to work. Thank you, Linda, for keeping me moving forward along the correct path. I would also like to thank the other members of my committee, Molly Duggan and Monty Sullivan, who provided support and contributed valuable additions, comments, and corrections to the study and to this dissertation. Your contributions helped strengthen my work and I hope that you are all satisfied with what we accomplished.

Other members of the ODU faculty who were inspirational that first summer were Mike Dingerson, Dennis Gregory, and Philip Reed. Had it not been for that intense, but stimulating, first summer institute I’m not sure that I would have continued with the program. However, your enthusiasm, friendliness, and encouragement inspired me to continue the journey through many personal life changes. Thank you all.

Thanks also to my mother who has always believed in me. Mom, I’m finally a doctor (but don’t ask me for a prescription)! To my daughter, Megan, I hope this will

inspire you to continue your education. Teenagers may know it all, but as we get older we have to go back to school to learn new things. Even a dad who knows all sorts of useless trivia can add to his knowledge base when he believes in life-long learning. I may have reached the end of this journey, but I've not stopped learning.

TABLE OF CONTENTS

	Page
LIST OF TABLES	xii
CHAPTER	
I. INTRODUCTION	1
SIGNIFICANCE OF THE STUDY	5
BACKGROUND	6
EVALUATION PLAN ELEMENTS	7
STUDENT EVALUATIONS	8
PEER EVALUATIONS	9
SELF-EVALUATIONS	10
ADMINISTRATIVE EVALUATIONS	10
RELATIONSHIP TO COMMUNITY COLLEGE LEADERSHIP	11
PROBLEM AND PURPOSE OF THE STUDY	12
RESEARCH QUESTIONS	16
OVERVIEW OF METHODOLOGY	16
THE STUDY POPULATION	16
THE STUDY DESIGN	17
DATA COLLECTION	18
DATA ANALYSIS	18
LIMITATIONS	19
II. REVIEW OF THE LITERATURE	21
INTRODUCTION	21
THE PURPOSE OF FACULTY EVALUATION	23
PERCEPTIONS OF FACULTY EVALUATION	23
EVALUATION TO IMPROVE FACULTY PERFORMANCE	24
FORMATIVE VERSUS SUMMATIVE EVALUATION	25
WHAT DUTIES SHOULD BE INCLUDED IN	
FACULTY EVALUATION?	28
DATA SOURCES FOR FACULTY EVALUATION	29
STUDENT EVALUATIONS	30
PEER EVALUATIONS	33
SELF-EVALUATIONS	35
ADMINISTRATIVE EVALUATIONS	38
FACULTY RESPONSIBILITIES SUBJECT TO EVALUATION	39
DEFINITION OF GOOD TEACHING	39
DUTIES TO INCLUDE IN THE EVALUATION PROCESS	41
LEVEL OF IMPORTANCE OF EACH DUTY	41
COMBINING VARIOUS DUTIES INTO AN OVERALL	
RATING	42

COMPREHENSIVE EVALUATION AS A CONTINUOUS IMPROVEMENT PROCESS.....	43
LIMITATIONS OF THE CURRENT EVALUATION SYSTEM.....	46
VARIATIONS IN THE VCCS EVALUATION SYSTEM.....	49
INFLUENCE OF THE LITERATURE ON THIS STUDY.....	50
III. METHODOLOGY.....	52
RESEARCH DESIGN AND QUESTIONS.....	52
VARIABLES.....	52
RESEARCH QUESTIONS.....	53
PARTICIPANTS.....	53
SURVEY POPULATION.....	53
MEASURE.....	55
SURVEY INSTRUMENT.....	55
THE SURVEY INSTRUMENT DESIGN.....	56
VALIDITY.....	58
RELIABILITY.....	58
PROCEDURE.....	59
ADMINISTRATION.....	59
FOLLOW-UP.....	60
DATA ANALYSIS.....	60
PRIMARY PURPOSE OF FACULTY EVALUATION.....	60
IV. RESULTS.....	63
QUANTITATIVE RESULTS.....	65
PURPOSE AND USE OF A FACULTY EVALUATION PLAN.....	65
FORMATIVE EVALUATION.....	69
SUMMATIVE EVALUATION.....	71
EVALUATION OF TEACHING ACTIVITIES.....	73
FACTOR ANALYSIS.....	81
EVALUATION OF TEACHING PERFORMANCE.....	81
EVALUATION OF SERVICE PERFORMANCE.....	84
PRIMARY ROLE.....	87
FACULTY DEMOGRAPHIC ANALYSIS.....	90
PRIMARY TEACHING AREA.....	90
YEARS OF FULL-TIME TEACHING.....	93
FACULTY GENDER DIFFERENCES.....	96
NO SIGNIFICANT INTERACTIONS.....	97
SUMMARY OF QUANTITATIVE RESULTS.....	98
QUALITATIVE ANALYSIS.....	100
STRENGTHS.....	101
LIMITATIONS.....	105
OTHER FEEDBACK.....	114
SUMMARY OF QUALITATIVE RESULTS.....	119
V. DISCUSSION.....	120
QUANTITATIVE RESULTS.....	120

THE PURPOSE OF FACULTY EVALUATION.....	120
COMPOSITION OF THE FACULTY EVALUATION PLAN.....	123
DATA SOURCES USED IN DETERMINING THE RATING.....	124
AREAS OF FACULTY RESPONSIBILITY FACTORED INTO THE FACULTY RATING.....	128
DEMOGRAPHIC DIFFERENCES IN RESPONSES.....	129
QUALITATIVE RESULTS.....	131
STRENGTHS OF THE VCCS PROCESS.....	131
LIMITATIONS OF THE VCCS PROCESS.....	132
SUGGESTED CHANGES TO THE VCCS PROCESS.....	133
LIMITATIONS.....	134
DIRECTIONS FOR FURTHER STUDY.....	136
SURVEY INSTRUMENT AS A TOOL FOR FUTURE RESEARCH.....	137
IMPLICATIONS FOR PRACTICE.....	137
VCCS POLICY REVISIONS.....	138
COLLEGE REVIEW.....	140
PROFESSIONAL DEVELOPMENT OPPORTUNITIES.....	141
CONCLUSIONS.....	141
REFERENCES.....	144
APPENDIXES	
A. MAIL CORRESPONDENCE TO COLLEGE PRESIDENTS.....	156
B. EMAIL CORRESPONDENCE TO COLLEGE ACADEMIC DEANS.....	157
C. ON-LINE FACULTY EVALUATION SURVEY.....	159
D. GIFT CARD LOTTERY FOR SURVEY PARTICIPANTS.....	170
E. <i>VCCS POLICY MANUAL SECTIONS 3.5 AND 3.6</i>	171
VITA.....	177

LIST OF TABLES

Table	Page
1. Faculty Member Concerns Related to Student Evaluation of Teaching.....	31
2. VCCS College Demographics for 2007-2008.....	54
3. Blueprint of Faculty Evaluation Survey Instrument.....	57
4. VCCS College Demographics for 2007-2008 and Number of Survey Respondents...	64
5. Purpose and Use of an Ideal Faculty Evaluation Plan: Overall Means.....	66
6. Purpose and Use of an Ideal Faculty Evaluation Plan:	
Faculty vs. Administrator ANOVA.....	67
7. Purpose and Use of an Ideal Faculty Evaluation Plan: Mean Comparisons	68
8. Overall Formative Faculty Performance: Overall Means and Standard Deviations...	69
9. Overall Formative Faculty Performance: Faculty vs. Administrator ANOVA.....	70
10. Overall Formative Faculty Performance: Mean Comparisons.....	71
11. Overall Summative Faculty Performance: Overall Means and Standard Deviations...	71
12. Overall Summative Faculty Performance: Faculty vs. Administrator ANOVA.....	72
13. Overall Summative Faculty Performance: Mean Comparisons.....	73
14. Teaching Performance Activities: Overall Means and Standard Deviations	73
15. Teaching Performance Activities: Faculty vs. Administrator ANOVA.....	75
16. Teaching Performance Activities: Mean Comparisons.....	76
17. Contributing Sources for Teaching: Overall Means and Standard Deviations.....	77
18. Contributing Sources for Faculty vs. Administrator Mean Comparison, ANOVA.....	78
19. Contributing Sources for Teaching: Mean Comparisons.....	79
20. Summary Wilks' λ Results for all Faculty vs. Administrator MANOVAs.....	80

21. Pattern Matrix Principal Component Analysis for	
“Evaluation of Teaching Performance” Items.....	82
22. Pattern Matrix Principal Component Analysis for	
“Evaluation of Service Performance” Items.....	85
23. Primary Area of Responsibility: Overall Means and Standard Deviations.....	88
24. Primary Area of Responsibility: Between-Subjects Effects, ANOVA.....	89
25. Primary Area of Responsibility: Mean Comparisons.....	90
26. Primary Teaching Area: Overall Means and Standard Deviations.....	91
27. Primary Teaching Area: Between-Subjects Effects, ANOVA.....	92
28. Primary Teaching Area: Mean Comparisons.....	92
29. Years of Full-Time Teaching Experience: Between-Subjects Effects, ANOVA.....	94
30. Years of Full-Time Teaching Experience: Mean Comparisons.....	95
31. Scheffe Post Hoc Analysis of Years of Full-Time Teaching Experience.....	95
32. Gender: Between-Subjects Effects, ANOVA.....	97
33. Summary Wilks’ λ Results for all Demographic MANOVA Calculations.....	97
34. Response Frequencies and Percentages for Emergent Themes for the Strengths	
of College Faculty Evaluation Plans.....	102
35. Response Frequencies and Percentages for Emergent Themes for the Limitations	
of College Faculty Evaluation Plans.....	105
36. Response Frequencies and Percentages for Emergent Themes for the Suggested	
Changes Needed for College Faculty Evaluation Plans.....	110
37. Response Frequencies and Percentages for Emergent Themes for Other Areas	
of Concern Related to the College Faculty Evaluation Plans.....	114

CHAPTER 1

INTRODUCTION

Faculty evaluation is a contentious process for both college teaching faculty members and their administrative supervisors for a variety of reasons. The primary reason for this contention is that the process typically tries to serve two conflicting purposes: (1) to provide constructive feedback to faculty members for the purpose of improving their teaching performance (formative evaluation) and (2) to evaluate faculty performance for purposes of promotion, tenure, reappointment, and/or salary adjustments (summative evaluation). According to Redmon (1999), faculty members and college administrators often have different perceptions about why an evaluation process is conducted. This difference of perception is just one symptom of a general lack of understanding of the importance of evaluation in education, and how that evaluation process should be conducted.

There is a fundamental distinction between measurement, rating, and evaluation. Measurement is the systematic process of assigning numbers or values to individual objects in a set to indicate differences between the objects and the magnitude of that difference (Arreola, 2007). Rating is a specialized type of measurement that involves the subjective process of selecting one number from a limited range of choices on a scale (as in a survey instrument) to assign a value to whatever is being measured or assessed. Evaluation is the process of interpreting measurements or rating data in order to make a value judgment on the degree to which the object under evaluation represents a desired quality (Arreola). According to Fitzpatrick, Sanders, and Worthen (2004), the primary purpose of evaluation is to arrive at a judgment of the worth of whatever is being

evaluated. In the case of faculty evaluation, it is the faculty members' performance of expected duties that should be evaluated. While there are many different ways of describing the process of evaluation, Fitzpatrick, Sanders, and Worthen emphasized that evaluation encompasses:

inquiry and judgment methods, including (1) determining standards for judging quality and deciding whether those standards should be relative or absolute, (2) collecting relevant information, and (3) applying the standards to determine value, quality, utility, effectiveness, or significance (p. 5).

Perhaps it is because the evaluation process leads to a statement about their worth as individuals, teaching faculty members have produced a wealth of literature on the topic. It is worth noting that the vast majority of the research has been conducted by faculty members who are primarily researchers at senior, four-year institutions yet nearly one-half of all undergraduate students in the United States of America are taught by community college teaching faculty members (AACC, 2009) who spend the majority of their time teaching and interacting directly with students – often teaching four to six classes per term. Other academic pursuits such as conducting scholarly research, publishing and grant writing, as well as engaging in service activities, play a minor (if not insignificant) role in the expectations for community college faculty performance. Therefore, many of the issues and concerns addressed in the faculty evaluation literature may not be entirely relevant to community college teaching faculty members.

Other reasons why the faculty evaluation process is so frustrating to the participants (students, faculty, and administrators) are (a) the widespread belief that the faculty evaluation process does not lead to any change or improvement, (b) performance

evaluation processes rarely, if ever, actually work the way they were intended, (c) the faculty evaluation summary rating often serves as the high stakes, sole determinant of a faculty member's opportunity for promotion and merit pay raises, and (d) most faculty have received little or no training for many of the tasks upon which their evaluation is based (Arreola, 2007). In fact, Arreola (2007) describes college teaching faculty members as "metaprofessionals" who are often evaluated based on their performance on tasks for which they have received no training or support such as: instructional research, design, delivery and assessment; course management; learning theory; information technology; technical writing; conflict management; budget development; and public speaking. Often that lack of training or support continues throughout the faculty member's career. College teaching is very much a "trial and error" occupation punctuated by annual performance evaluations and no intervening training or support. Ironically, Milton and Shoben (1968) noted that college teaching is one of the only professions in the world where people with no specific training are hired to perform a complex task.

Contributing to the contention inherent in faculty evaluation processes is the lack of agreement on (a) the evaluation elements (student, peer, self, administrative, etc.) that should be utilized by administrators when conducting faculty evaluations, (b) the reliability and validity of these elements, and (c) the relative importance placed upon each of these elements. Each of these evaluation elements will be discussed in detail in Chapter 2 of this dissertation. In addition, attempts to link student outcomes to teaching faculty performance cannot account for the wide range of student variables that influence performance in the classroom; thus making inferences about teaching quality based on student performance a risky practice (Fenwick, 2001).

With all of these potential areas of contention in any faculty evaluation process, it is not surprising that teaching faculty members and administrators in the Virginia Community College System (VCCS) have repeatedly complained to their Chancellor about how the process operates at their institutions. In response to criticism of the faculty evaluation process, the Chancellor included a goal in his *A Strategic Direction: Dateline 2009* that stated “Full-time teaching, professional and administrative faculty will be evaluated using a standard evaluation system approved by the Chancellor” (VCCS, 2003, p. 4). In keeping with his interest in the faculty evaluation process, the Chancellor formed the Faculty Evaluation Task Force in 2007 to review the status of college faculty evaluation plans in order to facilitate the creation of an improved faculty evaluation process for the colleges. To date, there have been no changes to the VCCS faculty evaluation process or policies as a result of the task force recommendations.

It must be clearly stated and understood that this study pertains only to the evaluation process for teaching (not professional or administrative) faculty members. In the VCCS, counselors and librarians hold the rank of professional faculty; and all administrators, with the exception of the college president, also hold faculty rank – yet teaching is not their primary duty. Therefore, this study is limited to consideration of the evaluation process for teaching faculty members only. Professional and administrative faculty members deserve to have similar consideration applied to their separate evaluation processes but that is beyond the scope of this study.

The work of the Chancellor’s task force is intended to address the concerns and limitations of the present teaching faculty evaluation system in Virginia’s community colleges. In particular, the task force was charged with reviewing current college faculty

evaluation plans for their compliance with VCCS policies, and to make recommendations on policy language that could be revised to improve the faculty evaluation process. In a related effort, this study focused on the evaluation process for teaching faculty in the VCCS and collected and analyzed administrator and teaching faculty member perspectives on the construction of an ideal faculty evaluation plan, and comments on how their own college plans could be improved. The recommendations of the task force will be combined with the results of this study to help shape revisions to the VCCS policy on teaching faculty evaluations, and college personnel may also use results from the study to help in the revision of their own faculty evaluation plans.

Significance of the Study

While much research has been conducted at four-year institutions, the amount of information available on faculty evaluation processes at two-year community colleges is scant at best. This is an interesting phenomenon when considering the number of students enrolled at community colleges in comparison to enrollments at four-year institutions. According to the American Association of Community Colleges, community college students make up about 43% of all undergraduates enrolled in higher education in the United States (AACC, 2009). In Virginia, community college students consistently comprise 50-60% of Virginia's undergraduate population (State Council of Higher Education for Virginia, 2010; VCCS, 2010b).

In contrast to senior institutions, all VCCS students are taught by faculty members (whether full-time or part-time) not by unsupervised graduate students (VCCS, 2010a). In addition, instruction of lower division undergraduate students is the primary activity required of community college faculty members while other scholarly pursuits such as

research and service only contribute a minor (if any) amount to a faculty member's annual evaluation review. One hundred percent of instruction to the majority of Virginia's undergraduate population is conducted by community college faculty members, therefore, community colleges need to have a consistent way of evaluating the quality of that instruction. A system-wide comprehensive faculty evaluation system will provide this needed consistent framework for evaluating faculty performance – primarily in the area of teaching.

Another glaring gap in the literature is the lack of assessment of administrator and/or faculty perceptions on faculty evaluation based on demographic factors such as the primary content area of instruction, gender, years of experience, and primary instructional (or supervisory) role. To be specific, there have been relatively few faculty evaluation studies comparing the perceptions of administrators and faculty members who are primarily associated with career and technical education, general education for transfer students, and/or developmental education. Although a few studies compare student perceptions of faculty members based on either the students' or the faculty members' gender, there is a lack of good data on how the faculty members and their administrators perceive the evaluation process when compared by gender. In a similar fashion, most investigators do not consider the faculty member or administrator's years of experience when conducting their studies. These demographic considerations will be examined and reported in this study as another contribution to the existing faculty evaluation literature.

Background

There are almost as many different ways to conduct a faculty evaluation as there are educational institutions across the nation. In concert with the rest of the country,

-

Virginia's community colleges are highly variable with respect to their faculty evaluation processes. In fact, each college in the VCCS has its own, unique, faculty evaluation plan that was developed from the same basic guidelines contained in the *VCCS Policy Manual* (VCCS, 2010a). The four major components of most faculty evaluation plans are (a) student evaluations, (b) peer evaluations, (c) self evaluations and (d) administrative evaluations. While faculty evaluation plans may be reduced to these four basic elements (student, peer, self, administrative), the manner in which some or all of these elements are combined to produce the final result is almost infinitely variable. These four basic elements of the teaching evaluation plan will be introduced here, but examined in more detail in Chapter 2 of this dissertation.

Evaluation Plan Elements

Since teaching is the primary job responsibility of community college faculty members, evaluation focuses on the assessment of classroom and/or on-line teaching of undergraduate students. Traditional faculty duties at senior institutions such as scholarship (research and publications) and service (to the department, institution, community or discipline) are not included this section since these duties are not commonly part of the expectation for community college teaching faculty members. While this does not exclude community college faculty members from engaging in such activities, and doing so may enhance the evaluation rating of some faculty members, a community college teaching faculty member may earn the highest level of evaluation rating without participating in any external scholarship or service activities. It is not uncommon for full-time community college teaching faculty members to expend all of their efforts on "internal" scholarship activities that relate directly to preparing for

classroom and laboratory instruction, and involvement in activities that enhance the success of students in their own classes.

Student evaluations. The bulk of the faculty evaluation literature concentrates on student evaluations of teaching performance. Although this concentration seems to be reasonable since students are the primary source of data regarding the performance of a faculty member in the classroom and/or laboratory, the use of student evaluations to evaluate faculty is not without controversy. Many faculty members oppose being “evaluated” by students for a variety of reasons that will be explored in more detail below; however, ample evidence supports student evaluations of instruction as reliable and valid indices of teaching quality. Centra (1993), Cashin (1999), and Arreola (2005) all found that, overall, student evaluations were reliable, stable and valid. Students tended to rate the same instructor similarly regardless of the subject matter. Conversely, different instructors teaching the same course received dissimilar ratings. Arreola (2007) cited several studies that showed high correlations between student evaluations of an instructor and the faculty member’s self rating, the administrator’s rating, and the ratings of colleagues.

In contrast, other findings highlight limitations associated with student evaluation of instruction (Addison, Best, & Warrington, 2006; Baldwin & Blattner, 2003; Davidovich & Soen, 2006; Foote, Harmon, & Mayo, 2003). Various student factors such as student motivation (Germain & Scandura, 2005), level of preparation (Langbein, 1994), grade point average (Millea & Grimes, 2002), gender (Ory, 2001), grade expectation (Yunker & Yunker, 2003), the personality of the faculty member (Russell & Gadberry, 2000), and the level of the course being taught/evaluated (Theall, 2005) may

have an influence on student evaluations of teaching. Therefore, student evaluations should be used with caution.

The inclusion of student evaluations of teaching is not meant to imply that students should also have a voice in evaluating other areas of faculty responsibility such as scholarly or service activities. Even though a college student's ability, biases, and/or mastery of course material may influence their evaluations, students remain the primary source of information regarding the performance of a faculty member in the classroom and/or laboratory, and student feedback should be included as one (not the only) component of a faculty evaluation plan – particularly for community college faculty members whose primary responsibility is classroom instruction (Arreola, 2007; Centra, 1993; Seldin, 1999).

Peer evaluations. Less well researched than student evaluations are peer evaluations of teaching performance, although the incidence of peer evaluation is growing (Osborne, 1998). As with student evaluations of teaching, peer evaluation has its supporters and detractors. According to Seldin (1999), for a peer evaluation to be truly meaningful, the peer evaluator needs to receive training on how to evaluate a fellow faculty member and must spend time reviewing course materials as well as visiting one or more class sessions of the person being evaluated. This is very time consuming and requires a high level of commitment on the part of the peer evaluator and the institution if the process is to result in any useful, reliable information. In addition, faculty members are hesitant to evaluate each other for a number of reasons, such as not wanting to pass judgment on a colleague, not wanting to spend the time and effort required to provide meaningful feedback, or fear that they will be judged harshly if evaluated by their peers.

Consequently, peer review will continue to remain a highly variable component of faculty evaluation plans unless colleges can devise a feasible procedure that is not overburdensome in terms of time invested in the processes.

Self-evaluations. Faculty members are often expected to provide documentation of their accomplishments over the course of the evaluation cycle. While self-evaluation may take the form of a simple checklist or a narrative summary, there has recently been an increased emphasis on the use of faculty portfolios for conducting both formative and summative evaluations of faculty. Melland and Volden (1996) described a portfolio as an organized compilation of materials that are carefully selected to be reflective of the range of a faculty member's teaching activities. They also recommended that portfolios contain three types of information (1) evidence of student learning, (2) the faculty member's philosophy of teaching, and (3) evaluative material from other sources, such as student and peer evaluations. In summary, the portfolio content should be an orderly and representative collection of the work that the faculty member and her students performed over the course of the evaluation cycle, which gives the reviewer a well-rounded view of the faculty member's activities, particularly those activities of which the reviewer would otherwise be unaware (Melland & Volden, 1996; Sain, 2008).

Administrative evaluations. Scant research explores administrative evaluation of teaching as a separate component of the faculty evaluation process. Seldin's (1999) survey of administrators at four-year liberal arts colleges on the subject of faculty evaluation in 1978, 1988 and 1999 found that college administrators rarely conducted personal classroom visits to obtain first-hand information on a faculty member's performance in the classroom. Instead, the administrators usually relied on second-hand

sources of information when evaluating teaching performance. According to Seldin, the most common sources of information used by administrators when completing faculty evaluations were student evaluations, evaluation by other administrators, and self-evaluation. Therefore, administrative evaluation tends to be a global summary of all other sources of input from first-hand sources to arrive at the overall annual performance evaluation rating that is used for both summative and formative purposes.

Relationship to Community College Leadership

Faculty evaluation in the VCCS is conducted primarily by the immediate supervisors of faculty, who are usually division deans or assistant deans and the first level of college academic administration. Academic deans are in arguably one of the most difficult positions at a college – between the faculty and the higher administrative levels. Deans must not only serve the needs of their teaching faculty members, but they must also ensure that those same faculty members follow college and system policies (McArthur, 2002). To effectively enforce policy, deans need to have good working relationships with their faculty members. According to McArthur, shared authority has many advantages. The culture of the institution or department, as well as the work ethic, attitude, and morale of the faculty, can be dependent on feelings of ownership and commitment. Administrative decisions are more readily accepted when faculty members have a voice in the decision-making process.

Faculty grievances against supervisors in the VCCS often depend on interpretations of policy. According to Section 3.13.0.a of the *VCCS Policy Manual* (VCCS, 2010a), if either the faculty member or her supervisor is found to be in violation of policy, that individual will lose the grievance. Since VCCS faculty members may

grieve their performance evaluations, the faculty evaluation process must avoid ambiguity and inconsistency. According to C. D. Lee, Associate Vice Chancellor of Human Services for the VCCS (personal communication, March 20, 2008), between 2005 and 2007, 83% of the VCCS faculty members received an “excellent” evaluation rating with another 15% receiving a “very good.” Such high rankings indicate that the current evaluation processes at Virginia’s community colleges lack objectivity, leading to the conclusion that supervisors feel that there is no need for improvement by the vast majority of their faculty. That is, VCCS faculty evaluations effectively demonstrate a ceiling effect. Instead of the faculty evaluation process serving as an objective, formative performance evaluation, the process has become almost entirely a high stakes summative prerequisite for faculty promotion and/or pay increase – or to serve as the springboard for a grievance when promotion and/or pay increase is held back due to a less-than-excellent evaluation rating.

Both faculty members and their supervisors need to agree that the faculty evaluation process is relatively fair and objective. The process should be centered on providing timely, constructive feedback on faculty member performance with the goal of encouraging continuous improvement without unfairly hindering faculty promotion and/or pay raise opportunities. Professional development opportunities should then be focused on continuous improvement of faculty member performance which benefits the faculty members, the students, and the entire college community.

Problem and Purpose of the Study

Faculty evaluation is a contentious area for both college faculty members and their administrative evaluators. This is primarily due to the fact that, according to *VCCS*

Policy Manual Section 3.6.1 (VCCS, 2010a) the process is used for two conflicting purposes: (a) formative evaluation to provide constructive feedback to faculty members for the purpose of improving their teaching performance, and (b) summative evaluation to evaluate faculty performance for purposes of promotion, tenure, reappointment, and/or salary adjustments. Virginia's community colleges therefore need a structured framework for a comprehensive faculty evaluation in order to meet both needs (formative and summative) and to provide a more consistent process across the state (Summers, 2007). Bringing consistency and a level of objectivity to this process should help to remove a great deal of the frustration and anxiety related to the faculty evaluation process currently felt by both faculty members and their administrative evaluators. In addition, Summers pointed out one factor that contributes to the lack of consistency in the VCCS faculty evaluation policy; there is not a strong connection between section 3.5 - Faculty Responsibilities and section 3.6 - Faculty Evaluations, see Appendix E. More consistent policy language may help provide a stronger framework on which the colleges could construct better evaluation plans.

To bring some harmony into the process of faculty evaluation, two conditions must be met. The first condition is a high degree of agreement between faculty members and their supervisors on exactly what aspects (metrics) of faculty performance are being evaluated, including explicit agreement upon the measurable standards that are expected for each rating level for those performance measures. The second condition is a common understanding that the evaluation plan can serve as a means for both measuring performance and providing the basis for developing individual faculty plans for continuous improvement (Arreola, 2007).

The purpose of this study is to address an area of critical importance to both VCCS faculty members and their supervisors. The faculty evaluation process remains a controversial issue in the VCCS (Minutes of the Advisory Council of Presidents, VCCS, 2008a). Faculty members contend that either (a) undeserving, under-performing faculty members have received the same excellent rating as they have earned themselves or (b) they have been arbitrarily given a rating at a level that is less than deserved due to a capricious supervisor. Deans lament they are afraid to assign a rating that is below excellent for fear they will be dragged into a lengthy and acrimonious grievance procedure for assigning an objective (often meritorious) rating, such as good or very good, to a faculty member. The vice presidents assert their faculty and deans are unhappy with the faculty evaluation process, yet they do not want the VCCS system office to devise a standard plan to be used at all of the colleges in an identical fashion; the vice presidents want to retain local authority to govern themselves. The presidents definitely want to retain their local authority to create and administer the faculty evaluation process but they also are in favor of improving that process (VCCS, 2008a).

This study utilized a survey instrument to identify common elements that could be used by the colleges when they reconsider their current faculty evaluation plans and to collect suggestions from administrators and faculty members that could possibly lead to improvements in VCCS policy language regarding faculty evaluations. The survey was based on Seldin's (1984) instrument and was modified to address the particular concerns of Virginia's community college teaching faculty and administrators. In addition this study addressed two salient gaps in the literature on faculty evaluation. First, although a plethora of studies addressed the evaluation of faculty at four-year institutions, little

research has been conducted at community colleges. Community college faculty are almost exclusively dedicated to direct interaction with students through teaching and advising, therefore, research focused on faculty at senior residential institutions who also conduct research and/or engage in significant activities that do not directly involve interactions with students is not appropriate. The student populations at senior residential institutions are also different from typical community college student populations, and comparing the involvement of the two groups of students in the faculty evaluation process is not necessarily valid. Therefore, studies based on senior residential students and faculty members do not necessarily apply equally to community colleges.

Second, little research compares faculty and administrator views on the faculty evaluation process based on demographic characteristics such as primary area of instruction or supervisory responsibility, years of experience, or gender. If demographic differences existed in faculty and/or administrator perceptions, this could provide college personnel with better information for revising their current faculty evaluation plans. For instance, if junior faculty members valued student feedback the most, perhaps that portion of their evaluation plan should be more heavily weighted than for senior faculty members, or vice versa. Career and technical faculty members might have preferred to have laboratory instruction as a major component of their evaluations while transfer faculty members did not. Males and females may have had very different preferences with respect to involvement in student and/or community activities. Without specifically analyzing the data for these types of demographic differences, the grouped data might have very well masked significant differences through a leveling out effect. Therefore, the following research questions were addressed.

Research Questions

1. What do the VCCS faculty members and administrators perceive to be the primary purposes of faculty evaluation (i.e. formative vs. summative)?
2. What do VCCS faculty members and administrators perceive to be the most important elements of a comprehensive faculty evaluation plan?
3. Do perceptions of the most important components used in the evaluation process differ as a function of demographic and background variables (i.e. primary role, primary content area, years of experience, and/or gender)?
4. What do VCCS faculty members and administrators perceive to be the strengths and limitations of the current faculty evaluation process?
5. What do VCCS faculty members and administrators suggest should be changed in the current faculty evaluation process?

Overview of Methodology

The Study Population

The potential study population included all academic deans and teaching faculty members from all 23 colleges in the VCCS as of spring semester 2009. There were approximately 95 academic deans and 2200 full-time teaching faculty members in the VCCS (VCCS, 2008c) at that time. The actual study population did not include administrators or full-time teaching faculty members from some of the colleges since two college presidents did not agree to have their staff members contacted for participation in the study.

The Study Design

The study employed survey methodology and addressed both descriptive and comparative research questions. The literature review formed the foundation for the survey instrument with selected questions modified from Peter Seldin's 1983 survey instrument (Seldin, 1984), Raoul Arreola's (2007) "partial list of possible faculty roles," and additional questions developed by the investigator. The survey instrument was divided into the following seven parts:

1. demographic information,
2. purpose and use of an ideal faculty evaluation plan,
3. formative evaluation of overall faculty performance,
4. summative evaluation of overall faculty performance,
5. evaluation of teaching performance,
6. evaluation of service performance, and
7. summary comments on the faculty evaluation process.

Close-ended rating scale items were primarily used in conjunction with a small number of open-ended questions to provide more in-depth description and responses. Data was collected through on-line survey questionnaires. The questionnaires consisted primarily of two scales based on rating scale items plus a smaller number of open-ended and demographic questions.

The first scale used was the "agreement" scale that asked the survey participants to rate their level of agreement to a series of questions concerning the purpose and use of a faculty evaluation plan. This four-point scale utilized response options of "strongly disagree," "disagree," "agree," and "strongly agree." The second scale used was the

“importance” scale that asked the survey participants to provide a response that indicated the relative importance that each of the various faculty roles should be given as a component in the overall faculty evaluation process. The four-point scale utilized response options of “not important,” “somewhat important,” “important,” and “very important.” This same scale was also used to ask the survey participants to rate the relative importance of including specific aspects of teaching performance, and service performance in the evaluation process.

Reliability and validity of the questionnaire were estimated. Experts in the field of performance evaluation and survey methodology were asked to review the survey for content validity and revisions were made based on their feedback. In addition, factor analysis was performed to verify construct validity. Cronbach’s Alpha was also utilized to assess inter-item reliability for each scale.

Data Collection

A letter requesting permission to contact college deans and teaching faculty members was sent to each college president (Appendix A). Upon receiving presidential permission, an invitation to participate in the study was e-mailed to all VCCS academic deans at that college (Appendix B). The invitation to participate contained a link to the on-line survey instrument. In addition to completing the survey, deans were asked to forward via e-mail the invitation to participate to their teaching faculty members.

Data Analysis

The data analysis was both descriptive and comparative. Means and standard deviations were calculated for scales and items within scales. Multivariate analysis of variance (MANOVA) calculations were performed to determine whether scale scores

differed as a function of role (faculty or administrator) and other demographic variables. Content analysis was employed to analyze responses to open-ended items.

Limitations

While the survey population included the academic deans and full-time teaching faculty members in the VCCS, only individuals who had an interest were likely to take the time to complete the survey. Therefore, the study population consisted entirely of self-selected individuals. However, due to the wide-spread interest in the topic of faculty evaluation within the VCCS, there were enough responses to allow meaningful data analysis. This study involved only full-time Virginia community college teaching faculty members and the administrators who evaluate them. Therefore, the results may not confidently be generalized to be representative of any other faculty members inside of Virginia or any other states. In addition, this study was not intended to address the concerns of professional faculty members on their evaluation process.

Reliability and validity of the survey instrument were not established prior to the study. In order to estimate reliability of the survey instrument, Cronbach's alpha was computed for each scale. Validity was assessed through the development of the survey blueprint, expert review, and factor analysis of the response data.

Self-report is a separate limitation associated with survey methodology. In addition, since participation in this survey was voluntary with no tangible reward for most participants, one must be aware that individuals who chose to respond tended to have strong reasons for responding; either in favor of, or in opposition to, the topic of the survey. Faculty evaluation is a controversial topic, so some strong opinions were

expressed. Assurances of maintaining respondent anonymity improved response candor for the participants.

CHAPTER 2

REVIEW OF THE LITERATURE

Due to the tremendous amount of literature available on the topic of faculty evaluation this review made no attempt to be comprehensive. Instead, the review was focused on the following areas (a) the purpose of faculty evaluation, (b) the composition of typical faculty evaluation plans, (c) the data sources used in determining the faculty rating, (d) the areas of faculty responsibility that may be factored into the faculty rating, (e) faculty evaluation as a tool for continuous improvement, and (f) limitations of the Virginia Community College System (VCCS) faculty evaluation process. The purpose of this review was to provide background on the issues associated with the faculty evaluation process, and to illustrate the value of the study in terms of the potential contributions to the vast literature on this topic. This study expanded two areas of the faculty evaluation literature that are not well represented (a) information on community college faculty evaluations and (b) demographic comparisons of administrator and faculty perceptions of the faculty evaluation process.

Introduction

Faculty evaluation has been a source of concern since at least the late 1920's, beginning with H. Remmers in 1927, so it is certainly not a new topic (Blum, 1936; Remmers, 1930). Since the 1970's there has been a flood of research conducted on the topic and searching the Education Resources Information Center (ERIC) database using the terms "evaluation" and "college faculty" produces a list of over 5,800 articles. These 5,800+ articles represent the literature that is available at this single resource and ERIC only contains information dating back to 1966.

In recognition of this fact that the literature is so rich with respect to the faculty evaluation process, it is necessary to state that this chapter does not provide an exhaustive review of the literature. A diligent researcher can find literature to support or refute just about any aspect of the faculty evaluation process; therefore, only representative aspects of the literature that pertained to the purpose of this specific study were covered. Where many authors reported similar findings, only representative or more scholarly sources were cited, and the limited nature of this study precluded a review of every aspect of the faculty evaluation process.

It is worth noting that the vast majority of the research related to the faculty evaluation process has been conducted at senior, four-year institutions, yet nearly one-half (43%) of all U.S. undergraduate students are taught by community college faculty members (AACC, 2009). At senior research institutions, students are often taught by graduate teaching associates while full-time faculty members frequently devote much of their time to scholarly pursuits that do not directly involve teaching. In contrast, community college faculty members spend the majority of their time teaching and interacting directly with students; Virginia Community College System (VCCS) faculty members are required to teach 12-15 credits and 15-20 contact hours per semester (VCCS, 2010a). Therefore, the previous faculty evaluation research may not be entirely relevant to community college faculty members.

This study addressed two salient gaps in the research on faculty evaluation. First, although a plethora of studies address the evaluation of faculty at four-year institutions, little research has explored community college faculty evaluation yet roughly half of all college instruction occurs at community colleges. Second, although student evaluation of

faculty members has been researched for many demographic differences, almost no research compares faculty and administrator views on the faculty evaluation process based on their demographic characteristics such as gender, years of experience, or primary area of instruction/supervisory responsibility.

The Purpose of Faculty Evaluation

Evaluation of teaching faculty members is often a contentious process, primarily because a single evaluation process serves two often conflicting purposes. Is the faculty evaluation process intended to provide *formative* feedback to faculty members so that they can improve their performance, or is the goal of the process to provide *summative* feedback to assist administrators in making personnel decisions related to retention, promotion, or dismissal of teaching faculty members? Are these two purposes mutually exclusive or are they two sides of the same coin? Research suggests faculty members may reasonably expect the evaluation process to be formative in nature for their personal use, while administrators are more likely to use the results of the faculty evaluation process in a summative fashion to inform decisions related to personnel matters such as raises, renewal of contracts, or for promotion and tenure decisions (American Association of University Professors, 2006; Campion, Mason & Erdman, 2000; Centra, 1993; Morris, 1997; Seldin, 1999; Worcester, 1993).

Perceptions of faculty evaluation. This difference of perception often leads to administrative apathy and/or faculty resistance when an institution attempts to institute or revise a faculty evaluation plan (Arreola, 2007; Schaffner & MacKinnon, 2002). In addition to apathy, other words or phrases used by faculty members to describe their feelings about the faculty evaluation process include “flawed” (Glenn, 2007),

“frustration, anxiety, distrust, disquieting and dissension” (Schaffner & MacKinnon), “tension” and “anxiety” (Yao & Grady, 2006), “controversial” (McCaig, 2002), and “a circus” (Cutler, 2007).

In addition to conflicting interpretations of the primary purpose for conducting faculty evaluations, a second major reason that the faculty evaluation process is so frustrating to the participants (students, faculty and administrators alike) is due to the widespread belief that the process does not lead to any real or significant change or improvement. Students often believe that neither the faculty members nor their supervisors pay attention to student evaluations of faculty members (Sojka, Gupta & Deeter-Schmetz, 2002). Indeed, many faculty members ignore or discount student evaluations while at the same time admitting that the student evaluations they do receive are basically the same year after year (Sojka, Gupta & Deeter-Schmetz; Yao & Grady, 2006). Despite such attitudes, Centra (1993), Seldin (1999), and Arreola (2007) all stated that student evaluations remain the single-most useful direct indicator of faculty performance in the classroom; but they should not be the only source of information used in the faculty evaluation process. Students are in the unique position of having prolonged, repeated exposure to faculty teaching performance and are, therefore, in the best position to rate that performance (Scriven, 2005).

Faculty evaluation to improve faculty performance. Performance evaluation processes rarely, if ever, actually work the way they were intended (Lee, 2006). Performance evaluations are usually designed to measure and *rate* performance (summative evaluation), not to *improve* performance (formative evaluation), yet faculty evaluation plans are purportedly designed to serve both summative and formative

purposes (Lee). Unfortunately, formative comments tend to carry a negative/punitive connotation when an individual's performance is evaluated for summative rating purposes. Arreola (2007) and Lee agree that the main problem with most performance evaluation plans is that they are aimed at *rating* performance (for promotion, tenure, etc.) versus *appraising* performance with the aim of working toward continuous improvement and professional development/enhancement.

Formative versus summative evaluation. Are these two purposes mutually exclusive or are they complementary to each other? One community college district that is comparable in size to the VCCS is the Maricopa Community College District in Arizona with approximately 220,000 credit students per year (Maricopa, 2009). Unlike the VCCS, Maricopa explicitly excluded the formative faculty evaluation plan from being used as a tool in the summative administrative faculty review for making personnel decisions (Maricopa, 2007). The much smaller (8,000 credit students) Itawamba Community College in Mississippi (Brock, Chrestman, & Armstrong, 1998) also separated their formative faculty review process from their summative faculty evaluation process.

Despite these examples, it is more common that one faculty evaluation process is used for both formative and summative purposes. Morris (1997) suggested faculty improvement (formative evaluation) and institutional accountability (summative evaluation) were two goals met by a single faculty evaluation process. To conform to accreditation criteria, Texas community colleges, by law, have faculty evaluation plans that are both formative and summative (Campion, Mason, & Erdman, 2000). A properly constructed and conducted faculty evaluation process can meet these two goals, but this is

not a simple process: it requires a combined effort on the parts of both faculty members and their administrative supervisors (Arreola, 2007; Lee, 2006).

One might also ask if there are any other purposes for performing faculty evaluations. The underlying unspoken goal of the faculty evaluation process is improvement in the quality of student education (McGee, 1996; South Texas College, 2004). If indeed the underlying reasons for conducting the faculty evaluation process are improved teaching effectiveness and quality of student education, then using the results of the faculty evaluation process for both formative and summative purposes makes sense. A formative process provides faculty members with specific information from students, who are the most frequent observers of teaching (Scriven, 2005) and faculty members can use this information to improve their teaching performance. In turn, a summative process can provide administrators with more global information from students (Cashin & Downey, 1992), and other contributors, to assist in the decision processes of retention and promotion of “good” faculty or the dismissal of “not good” faculty members.

Research also suggests one reason for conducting faculty evaluation is to ensure institutional financial accountability (Fairweather, 2002; Koops & Winsor, 2006; McGee, 1996; Morris, 1997). Arreola (2007) stated that public demands for accountability and the resulting legislative mandates have led to the use of faculty evaluation processes as accountability performance measures. In this climate, it is important to remember that Virginia’s community colleges are public institutions that rely heavily on state and federal funds to operate. Currently, there are no legislated accountability measures in Virginia that tie institutional funding to faculty or student performance. However, further

reductions in state funding could lead legislators to consider such an option even though performance-based funding has not proven to be effective in other systems (Harbour & Jaquette, 2007).

Centra (1993) stated that even formative evaluation can only lead to improved teaching if four conditions are met: (a) the faculty member must gain new knowledge as a result of the process, (b) the faculty member must value the source of the evaluative information, (c) there must be a mechanism that allows and encourages the faculty member to change, and (d) there must be some sort of intrinsic or extrinsic motivation for the faculty member to change. If a college has an effective faculty evaluation process the end results should therefore be improved teaching and retention of the best teachers; thus leading to a high quality educational experience for the students (Centra). However, some of the negative reviews found in the literature of faculty evaluations and the faculty evaluation process clearly tell us that not many institutions of higher education have effective faculty evaluation processes in place (Redmon, 1999).

Developing a functionally useful comprehensive faculty evaluation plan involves the recognition that both formative and summative evaluation results must be achieved through a single process. Arreola (2007) acknowledged that most faculty evaluation systems can serve both formative and summative purposes and the key to developing a comprehensive faculty evaluation system avoids administrative apathy and faculty resistance. Administrative apathy is counteracted by gaining support of a high-ranking administrator and involving that person in the process. Faculty resistance is reduced through direct involvement of faculty members from the earliest phase of plan development.

What Faculty Duties should be Included in Evaluation?

Another reason that many faculty members view any faculty evaluation process as flawed is they have received little or no training for many of the tasks upon which they are evaluated (Arreola, 2007). Indeed, Milton and Shoben (1968) noted that college teaching is one of the only professions in the world where people with no specific training are hired to perform a complex task. Faculty members are typically hired based on their subject matter expertise and/or research skills with little regard for whether or not they are trained on how to teach, develop academic programs, construct and evaluate exams for students, etc. Arreola described college teaching faculty members as “metaprofessionals” or individuals expected to demonstrate a high level of expertise in several skill areas beyond their traditional areas of expertise. Teaching faculty members are often evaluated on their performance in areas where they have little training or support. Often that lack of training or support continues throughout a faculty member’s career. College teaching is typically a “trial and error” occupation punctuated by annual performance evaluations and no intervening training or support (Grubb, et al., 1999). Scriven (2005) stated that evaluation is often a punitive or threatening process that leads to resentment and suspicion where there is not a comprehensive approach to the process.

A faculty evaluation plan that is meaningful for both faculty members and their supervisors requires strong agreement on the data sources, component areas, and specific elements within these areas to include in the plan. While there is not universal agreement, the following major components appear in most faculty evaluation plans (a) student ratings, (b) peer evaluations, (c) self evaluations, and (d) supervisor evaluation.

Data Sources for Faculty Evaluation

In his 1996 survey of 247 community colleges in ten states, McGee found that the three most common sources of information utilized in faculty evaluation plans were student (92%), supervisor (84%), and self (72%) evaluations. In addition, classroom visits were also utilized in 79% of the cases but it was unclear whether peers, supervisors, or a combination were conducting the visits. In 2002, Paulsen suggested the three most common sources of data were student ratings, peer evaluations, and self-evaluations/portfolios. South Texas College revised their faculty evaluation plan process in 2004 and determined that their faculty evaluations would be composed of instructor self-evaluations (30%), classroom observation by supervisor or lead instructor (30%), student evaluations (20%), and a summary administrator review (20%). As each of these resources made clear, utilizing several sources of information to provide a balanced faculty evaluation plan is very important (Centra, 1993; Seldin, 1999; Theall, 2005). Arreola (2007) added it is important to specify the proportional weight each source of information has on the total evaluation.

Four major components of faculty evaluation plans (a) student evaluations, (b) peer evaluations, (c) self evaluations, and (d) administrative evaluations were reviewed. The literature is highly variable with respect to the value of each of these components of faculty evaluation, so each must be considered with the proviso that there is no universal agreement as to the validity or reliability of most evaluative instruments. The literature on faculty evaluation is so rich that one may find several articles to support nearly any position (pro or con) one wants to take on any aspect of the process, particularly with respect to student evaluations.

Student evaluations. Many institutions develop their own forms, questionnaires or checklists to be used in the faculty evaluation process; others use commercially developed forms, questionnaire instruments, or processes (Theall, 2005). The area of student evaluations is the only one of the four common components where there are several commercially-available evaluative instruments. While these instruments may have a high degree of reliability, the validity of some specific survey items is often questioned by faculty members (Baldwin & Blattner, 2003). Consequently, the bulk of the literature on faculty evaluation concentrates on student evaluations of teaching performance, and one may easily find articles that support both sides of almost any concern that faculty members have with student evaluations of teaching performance - with the apparent exception of the general agreement that student motivation/preparation level has an effect on student evaluations of teaching performance.

Representative examples of pertinent literature are summarized in Table 1, which delineates the myriad of factors influencing student evaluations and illuminates the complexity of interpreting these evaluations, which contributes to faculty members' concerns about their use in the faculty evaluation process. For each factor identified in the literature, one can easily locate at least one study that concurs and at least one study that refutes the import of these factors. The only issue that does not seem to elicit disagreement is student evaluations of faculty members are influenced by the level of student motivation and/or preparation. That is, students who are better prepared for a course and who are motivated to do well tend to give higher evaluation ratings to the course and instructor. Therefore, the variability of viewpoints is likely a reflection of wide-spread opinion of student evaluations within the realm of higher education.

Table 1

Faculty member concerns related to student evaluation of teaching.

Student Evaluations	Agree	Disagree
should be an integral part of the faculty evaluation process	Arreola, 2007; Cashin, 1999; Centra, 1993	Wolfer & Johnson, 2003
are valid	Arreola, 2005; Hobson & Talbot, 2001; Ory, 2001; Theall, 2005	Langbein, 1994; Yunker & Yunker, 2003
are reliable	Hobson & Talbot, 2001; Ory, 2001; Seldin, 1993; Theall, 2005; Yao & Grady, 2006	Wright, 2006
are influenced by expected course grades	Germain & Scandura, 2005; McPherson, 2006; Millea & Grimes, 2002; Ory, 2001; Yunker & Yunker, 2003	Blum, 1936; Centra, n.d.; Donnelly, 2006; Russell & Gadberry, 2000
are influenced by the personality of the faculty member	Germain & Scandura, 2005; Russell & Gadberry, 2000	Ory, 2001
are influenced by the gender of the faculty member	Basow & Silberg, 1987; Germain & Scandura, 2005	Foote, Harmon & Mayo, 2003; Ory, 2001; Theall, 2005
are influenced by the gender of the student	Germain & Scandura, 2005; Ory, 2001	Downey, 2003; Foote, Harmon & Mayo, 2003; Theall, 2005
are influenced by the level of the course being taught (lower division, graduate, etc.)	Ory, 2001; Theall, 2005; Yunker & Yunker, 2003	Langbein, 1994; Wolfer & Johnson, 2003
are influenced by the level of course difficulty	Addison, Best & Warrington, 2006; Birnbaum, 1999	Dee, 2007; Millea & Grimes, 2002
are influenced by class size	Glenn, 2007; McPherson, 2006; Theall, 2005	Centra, 1993
are influenced by student motivation/preparation level	Davidovitch & Soen, 2006; Germain & Scandura, 2005; Langbein, 1994; Ory, 2001	

Studies vary significantly in the type of course or institution being evaluated – introductory, general education courses versus applied, program specific courses; lower-division versus upper-division or graduate-level courses; theoretical versus applied or performance-based courses; open-door community college versus selective liberal arts college versus highly selective research university; etc.. As stated by Yunker and Yunker (2003), some students' negative attitudes toward the subject matter might contribute to negative attitudes toward their instructors. At the community college level, most general education courses are populated by students who are "less-than-enthusiastic" about the subject matter. The fact that most student rating scores are positive, the Student Instructional Report II (SIR II) comparative mean for 2-year institutions is roughly "very good," indicates that students are usually satisfied with their experiences in most classes, according to the Educational Testing Service (ETS, 2006b). Indeed, the comparative mean score for the SIR II question relating to the overall rating of teaching and learning for a course is 4.04 out of 5.0 (ETS).

Cashin and Downey (1992) determined that global, summative ratings were appropriate for administrative purposes, while individual formative ratings were appropriate for formative feedback. Cashin and Downey reasoned that students are assigned a single, final course grade to represent the multidimensional aspects of their learning in a course; therefore a similar method could be used for faculty evaluations. When an individual faculty member's student rating score is significantly below the comparative mean score for her discipline there is probably something going on in the class that deserves investigation (Cashin & Downey).

The ETS maintains a web site with research articles posted on the topic of student evaluations of teaching. Not surprisingly, the ETS states that with over 2,000 articles on student evaluations of teaching referenced in ERIC, the bulk of the studies have been favorable with respect to student evaluations, their reliability, and their validity (Centra, n.d.). Hobson and Talbot's (2001) review of the literature (much of it the same as cited in the ETS article by Centra) concluded different researchers varied immensely in their views on the topic of student evaluations of teaching. Hobson and Talbot stated "Despite discrepancies in opinions and research findings on the validity of student rating, it is essential for faculty to understand that student evaluations are – and probably will continue to be – the primary institutional measure of their teaching effectiveness" (2001, ¶7). In the accounting classes they studied Read, Rama, and Raghunandan determined an inverse relationship existed between the emphasis placed on teaching and the weight given student evaluations (2001).

Sojka, Gupta, and Deeter-Schmetz (2002) determined both faculty and students agreed that instructors did not regard faculty evaluation seriously. Even so, both groups agreed that student evaluations remain useful and should not be eliminated. Nuhfer (2005) cautioned that no single measure, such as student satisfaction ratings, could adequately capture or describe a complex activity such as teaching. Nuhfer went on to state that "student ratings alone cannot capture 'good teaching,' prove that learning occurred, or serve to show outcomes were met" (p. 14). Therefore, a comprehensive faculty evaluation system should not rely on student evaluations alone.

Peer evaluations. Peer evaluations are less well researched, but evidence suggests that this form of evaluation is growing (Osborne, 1998). As with student evaluations of

teaching, peer evaluation has its supporters and detractors. For a peer evaluation to be truly meaningful, the peer evaluator should receive training on how to evaluate a fellow faculty member and must spend time reviewing course materials as well as visiting one or more class sessions of the person under evaluation (Seldin, 1999). This is obviously very time consuming and requires a high level of commitment on the part of the peer evaluator and the institution to result in any useful, reliable information.

In 1994, the North Carolina General Assembly made peer evaluations of faculty members a mandatory part of their tenure, promotion and reappointment process (Yon, Burnap & Kohut, 2002). However, there were broad guidelines to follow for each department within a given institution – much like the guidelines provided for VCCS faculty evaluation plans. Therefore, there is no more than a cursory commonality between the peer evaluations used at different departments within the same school. Yon, Burnap and Kohut reviewed the peer evaluation plans from 12 out of 30 departments at the University of North Carolina, Charlotte, and found that 60% of the departments required at least two people to perform classroom observations. All departments followed a three-part process: (a) a pre-observation review of course materials and an interview with the instructor being reviewed; (b) the in-class observation; and (c) a post-observation review with the students and/or the faculty member.

Unlike most faculty members at senior institutions, community college faculty members teach an average of 12 to 15 credit hours (4-5 classes) each semester, allowing few opportunities for faculty to observe other faculty members' classes or engage in such a time-consuming process. Consequently, peer review will continue to remain a highly variable component of VCCS faculty evaluation plans unless colleges can devise a

feasible procedure that is not over-burdensome in terms of time invested in the processes and the time required to provide useful feedback to the faculty member and his supervisor. Yon, Burnap, and Kohut (2002) concluded that peer observation reports could play an important part in the evaluation of teaching. The authors agree that the peer evaluation process is still being developed, and both the faculty reviewers and those being reviewed are gaining sophistication with the types of questions they are asking each other.

Self-evaluations. Self-evaluation may take the form of a simple checklist, a narrative summary, or an extensive portfolio. While a checklist has the advantage of simplicity – both for the faculty member to complete and for the evaluator to assign a rating - there is no evidence to support the validity of a faculty member’s self-ratings. A narrative summary by the faculty member of her accomplishments over the evaluative period has the advantage of providing the faculty member an opportunity to remind her evaluator of all of the “extra” activities in which she participated that would not show up in a simple checklist. In addition, a narrative should be accompanied by supporting documentation to provide evidence of the year’s accomplishments. The presence or absence of such supporting documentation provides the evaluator with a sound basis for assigning a summary rating (Fleak, Romine & Gilchrist, 2003).

The most comprehensive form of self-evaluation is when the faculty member creates a portfolio. Portfolios have been used for conducting both formative and summative evaluation of faculty. Melland and Volden (1996) defined a portfolio as a “compilation of carefully selected materials reflective of one’s teaching activities, presented in an organized manner by an individual faculty member” (¶2). They also

recommended that portfolios contain three types of information: “1) material reflective of student learning; 2) material from the faculty member reflective of teaching practices; and 3) evaluative material from others such as students, colleagues, or alumni” (§3).

Appling, Naumann, and Berk (2001) separated a teaching portfolio into two parts: (a) reflective analysis and (b) artifacts. Reflective analysis includes a “narrative of the educator’s philosophy of teaching and learning and a description of teaching responsibilities and activities” including courses taught, number of students, guest presentations given, scholarly projects/papers, training, grants, etc. (p. 250). Artifacts are “evidence or documentation in appendices to support claims in the reflective analysis section” and may include course syllabi, copies of student exams, handouts, student evaluations, faculty evaluations of students, etc. (p. 250). Faculty may also include a videotape or computer disk of a lecture or presentation to support their reflective analysis.

Faculty members should be thoughtful about the content that they include when they construct their portfolios. For example, the student input portion should be more than reproductions of student evaluations that are otherwise available to the supervisor. Student evaluation material could contain explanatory information that helps to clarify very high or very low student evaluation ratings. If the faculty member includes additional questions that are not part of the standard student survey instrument, student responses to these questions should be presented in this section of the portfolio. Examples of student work (Baldwin & Blattner, 2003) and/or student letters of thanks or support might also be included in the portfolio.

The portfolio content should be an orderly and representative collection of the work that the faculty member performed over the course of the evaluation cycle, which

gives the reviewer a well-rounded view of the faculty member's activities – particularly those activities of which the reviewer would otherwise be unaware (Melland & Volden, 1996) or activities that clearly show the faculty member's efforts in the best light. While portfolios do offer teaching faculty members the opportunity to provide evidence of their professional activities that are not captured by student evaluation questionnaires, etc., they also may become so complex and comprehensive that they become impossible to evaluate (Theall, 2005).

For instance, how could the range of materials included in a portfolio be objectively evaluated against similar or dissimilar materials submitted by other faculty members? Against what rating standards could portfolio materials be gauged? What evidence of effectiveness is required for the material to be valid and acceptable for inclusion in the faculty evaluation process? Theall asked: "How, for example, should one's philosophy of teaching and learning be judged?" (p. 3) In short, with no "standards" for portfolio construction and evaluation, inclusion of a portfolio has questionable worth in an objective evaluation system other than to provide evidence of various accomplishments.

In a recent study of North Carolina Community Colleges that utilize faculty portfolios as a standard component of their faculty evaluation process, Sain (2008) found that administrators and faculty members differed in their view of the value of portfolios. Faculty members felt that the self-reflection upon their work was the most valuable part of assembling their portfolios, while administrators viewed the portfolios as a comprehensive source of documented faculty performance – student evaluations,

projects, syllabi, annual goals, etc. Both faculty and administrators expressed concerns over the amount of time involved in utilizing the portfolio process (Sain).

Administrative evaluations. Very little information was readily available in the literature concerning administrative evaluation as a separate component of the faculty evaluation process. Perhaps this is because faculty members who do conduct research on the faculty evaluation process feel that they have no control over (or will receive little cooperation from) their supervisors and, therefore, their studies will necessarily have very small sample sizes. In addition, faculty members would most likely not be given access to the evaluation ratings of their peers – so the amount of information that they could study would be minimal. Faculty members have much more control over student data, typically anonymous, and can easily conduct research related to student evaluations of teaching where sample size is not an issue. On the other hand, administrators may be less inclined to conduct research and publish papers on an activity that is quite often acrimonious and unpleasant to conduct, or where they fear being sued (McLean, 1994).

Seldin (1999) surveyed administrators at four-year liberal arts colleges on the subject of faculty evaluation in 1978, 1988 and 1999 and found that college administrators were widely divided on the use of personal classroom visits in order to obtain first-hand information on a faculty member's performance in the classroom – 40% of those surveyed “always used” classroom visits while another 31% said that they “never used” classroom visits (p. 17). Traditionally, the administrator usually relied on second-hand sources of information when evaluating teaching performance. The most common sources of information used by administrators when completing faculty

evaluations were: (a) student evaluations, (b) evaluation by the department chair, (c) evaluation by the dean, and (d) self-evaluation (Seldin, 1999).

Perhaps this administrative reliance on second-hand data fosters a sense of distrust between faculty and administrators. Redmon (1999) stated that community college faculty members, “generally share a belief that administrators should be more willing to share resources and power, allow for creative growth and development in teaching, and allow for greater adaptability in showcasing their professional growth” (p. 57). Also, Paulsen (2002) stated that “clarifying the expectations that institutions and departments have for their faculty and that faculty have for their own performance are central to a successful faculty evaluation system” (p.5). At many institutions, this sharing of information between administrators and faculty members is woefully lacking (Seldin, 1999; Worcester, 1993).

Faculty Responsibility Areas Subject to Evaluation

According to Centra (1993), Seldin (1999), and Arreola (2007), some of the major stumbling blocks that impede the development of a useful faculty evaluation plan, are the lack of agreement on: (a) a definition of good teaching, (b) the duties or activities in which a faculty member engages that should be included in the evaluation process, (c) the level of engagement in these activities that is worthy of a particular ranking, and (d) how these activity ratings should be combined into an overall evaluation rating (e.g. excellent, very good, good, fair, or poor).

Definition of good teaching. In the book, *Honored but Invisible: An Inside Look at Teaching in Community Colleges*, Grubb, et al. (1999) wrote about the “community college paradox.” The paradox is that community colleges pride themselves on being

focused on teaching students (as opposed to research, etc.), yet they provide very little attention to the development of their own faculty members who often work in relative isolation. When community colleges fail to assume the responsibility for providing professional development for improvement of instruction, such an approach reinforces the view that good teachers are “born, not made” (Grubb, et al.). If colleges truly want to refute the idea that good teachers cannot be developed and nurtured, they must first define “excellent teaching” so that it can be quantified for evaluative and professional development purposes (Arreola, 2007; Seldin, 1999).

While there are some obvious attributes common to individuals who are widely considered to be good teachers, there is enough difference of opinion to make this a difficult question to answer (Grubb, et al., 1999). What exactly are the characteristics of good teaching, and can one confidently state that good teaching leads to student learning? Chickering and Gamson (1987) delineated *Seven Principles for Good Practice in Undergraduate Education*. The seven principles that Chickering and Gamson said good teachers possessed were:

- (1) encourages contact between students and faculty; (2) develops reciprocity and cooperation among students; (3) uses active learning techniques; (4) gives prompt feedback; (5) emphasizes time on task; (6) communicated high expectations; and (7) respects diverse talents and ways of learning (p. 3).

Measuring how well these seven principles are followed by teaching faculty members is not easily accomplished. However, questions directly related to these principles are part of the SIR II survey instrument devised by the ETS and utilized nationally for collecting information from students about their classroom experiences

(ETS, 2006a). The SIR II survey instrument attempts to have students rate their educational experiences in eight areas: (1) course organization and planning, (2) communication, (3) faculty-student interactions, (4) assignments, exams and grading, (5) course difficulty and workload, (6) course outcomes, (7) student effort and involvement, and (8) supplementary instructional methods (ETS, 2006a).

Duties to include in the faculty evaluation process. In his study of 247 community colleges, McGee cited a 1989 Carnegie Foundation study on teaching in higher education, which reported that “while 77% of all professors in comprehensive universities considered teaching their primary focus, 93% of professors in two-year schools felt that teaching was their main occupation” (1996, p. 64). McGee’s own study found that the top five factors which contribute to the evaluation process were, in descending order: classroom teaching (88.8%), student advising (45.5%), campus committee work (40.1%), length of service (34.3%) and personal attributes (19.2%) (p. 35). Again, there is a heavy emphasis placed upon classroom teaching in community colleges, yet “most postsecondary instructors have no formal preparation in teaching methods” (Grubb, et al., 1999, p. 26).

Level of importance of each duty. Depending on the specific duties and responsibilities assigned to individual faculty members, the relative importance of those duties can be quite different from one faculty member to another within a given division, or from one division to another. According to Seldin (1999), any duties or responsibilities used as part of the faculty evaluation process must be “job-related and subject to empirical validation” (p. 223). However, as Centra (1993) stated: “Not everything that counts can be counted and not everything that can be counted counts” (p. 176).

Therefore, using multiple measures from a variety of sources is necessary for triangulating data to confirm the assignment of quantitative levels of importance to specific duties and responsibilities. Faculty members and their immediate supervisor should have the flexibility to negotiate particular weighting values that will be utilized to determine the relative importance of particular components of individualized performance plans. The potential negotiation range should be established at the department, division, or college level when the comprehensive college evaluation plan is developed (Arreola, 2007).

Combining the various duties into an overall evaluation rating. Regardless of which particular duties are utilized, and how they are weighted, there must be an agreed-upon value (or range of values) assigned to each evaluative element of the comprehensive plan. Each institution, or functional unit in the institution, has to determine the manner in which duties are selected for inclusion in the evaluation process, what weightings are assigned to each duty, and what weightings are assigned to each of the information sources that contribute to the overall evaluation rating (Arreola, 2007; Centra, 1993; Seldin, 1999).

For example, Itawamba Community College placed 75% of the faculty evaluation weighting on teaching, 10% on service to the college, 5% on service to the profession or industry, and the remaining 10% on professional development - including creative and/or scholarly production (Brock, Chrestman & Armstrong, 1998). The Itawamba breakdown was according to faculty duty areas, not according to the four *sources* of information - although student evaluations contributed 50% of the weight for the entire evaluation rating with the balance coming from a combination of self and administrator evaluation.

As mentioned earlier, South Texas College (2004) determined that their faculty evaluations would be composed of: instructor self-evaluations (30%), classroom observation by supervisor or lead instructor (30%), student evaluations (20%), and a summary administrator review (20%). Adding together the weighted scores from these four areas at each of these institutions leads to a composite overall annual rating value.

While an annual summative evaluation rating is useful to administrators for making personnel decisions, and the formative student evaluation feedback is useful to the faculty members for self-reflection; if the end result of the evaluation process is intended to lead to improved instruction and a better educational experience for students, there should be some specific improvement (or professional development) plan to follow the assignment of an annual performance rating (Arreola, 2007; Centra, 1993, Murray, 2002). This leads us to Scriven's comment that the "implication is that the evaluation of faculty performance must be linked with institutional programs that support professional development as a necessary element in improving overall institutional performance" (2005, p. 9).

Comprehensive Evaluation as a Continuous Improvement Process

The concept of continuous improvement of teaching faculty performance is not a new idea; it is actually built into the accreditation standards under which institutions of higher education are intended to operate. Virginia's community colleges are all regionally accredited by the Southern Association of Colleges and Schools – Commission on Colleges (SACS-COC or SACS) and, as such, are subject to the accreditation standards published by SACS in their *Principles of Accreditation* (2009), including the standards related to faculty evaluation and professional development. These standards read as

follows: “The institution regularly evaluates the effectiveness of each faculty member in accord with published criteria, regardless of contractual or tenured status,” and “The institution provides ongoing professional development of faculty as teachers, scholars, and practitioners” (p. 30). Although there is no explicit link between these two SACS principles other than their proximity to each other, they are the direct descendants of the previous SACS *Criteria for Accreditation* (1997) that stated: “...The institution must demonstrate that it uses the results of this evaluation for the improvement of faculty and its educational programs” (§4.8.10). Therefore, there is the expectation that VCCS institutions will conduct faculty evaluations and use the results for professional development of their faculty members.

If the goals of conducting faculty evaluations are truly to correct poor performance, sustain good performance, and improve all levels of faculty performance then, according to Lee (2006): “The performance management process should be future oriented and focused on information, feedback and description” (p. 13). Traditional performance evaluations and appraisals typically do not work to improve performance since they were not designed to do so; they were designed to measure and rate the performance with no thought toward improving future performance. In other words, evaluation “denotes an assessment of worth and quality, and it explicitly implies a focus on the past” (Lee, 2006, p. 23). Faculty performance is a dynamic, on-going and multidimensional activity that cannot be adequately reduced to a single instrument that is utilized on an annual basis. Therefore, it is important that a comprehensive faculty evaluation process should allow for multiple data sources and the collection of relevant

data throughout the evaluation cycle – even in institutions that conduct the process on an annual schedule (Lee, 2006).

According to Arreola (2007), the faculty evaluation process should revolve around a performance plan that is reviewed each year at the start of the evaluation cycle, and both the faculty member and his supervisor should agree upon performance goals for the coming year and how those goals will be measured. Different departments at the same school can have different versions of the “college” plan, and individual faculty members can negotiate their individual plans within the departmental framework (Arreola, 2007). Faculty evaluation plan development and execution is, by necessity, a collaborative venture from start to finish. When faculty and administrative members work together as a team to develop and implement a plan that has meaning for each of the participants, there is the increased likelihood that the plan will be accepted by the majority of both supervisors and their faculty members. A meaningful plan is based upon negotiated standards that are supported by data collection and analysis (Arreola; Centra, 1993; Seldin, 1999).

The American Association of University Professors (AAUP) stated that an institution’s commitment to teaching should be focused on “obtaining first-hand evidence of teaching competence” (AAUP, 2006, p. 202) and echoed Schaffner and MacKinnon’s (2002) view that the evaluation process should be combined with professional development efforts to assist faculty members in becoming better teachers. McGee stated in 1996 that continuous improvement was a reason for conducting faculty evaluations and Schaffner and MacKinnon stressed the importance of “building a climate of continuous quality improvement” (p. 3). Schaffner and MacKinnon (2002) and Scriven (2005) also

found that in the absence of alignment between faculty evaluation results and professional development opportunities, faculty members perceived the evaluation process to be punitive, which led to suspicion of administrative motives and further inhibited faculty improvement. The linkage between professional development and faculty evaluation was described by Scriven as “a necessary element in improving overall institutional performance” – a goal of the faculty evaluation process as discussed earlier in this chapter.

The Limitations of the Current VCCS Evaluation System

A significant flaw in the VCCS faculty evaluation process is the annual faculty evaluation rating serves as the high stakes, sole determinant of a faculty member’s opportunity for promotion and merit pay raises (Summers, 2007). Summers also found that the VCCS annual summative faculty rating had such a profound economic impact on faculty members that administrators were reluctant to assign faculty evaluation ratings below the top level of “excellent” for two related reasons: (1) they did not want to stand in the way of a faculty member’s opportunity to receive a raise or to be promoted when the faculty member was performing in a satisfactory (“good” or better) manner, and (2) a faculty member who received a less than excellent rating would likely file a grievance against her supervisor – even when she received a meritorious rating of “good” or “very good.”

The VCCS policy section regarding faculty qualifications and promotions requires that the promotion of a teaching faculty member to the level of associate professor or full professor requires two consecutive years of “excellent” evaluation ratings. In addition, merit pay raises must be graduated so that those who receive

excellent summary ratings also receive the largest pay increases (VCCS, 2010a). To put this into perspective, over the past ten academic years the faculty members in the VCCS received no raises for four years (2001-02, 2002-03, 2009-10, and 2010-11), and they received only two- to four-percent raises in four of the other years (VCCS, 2009; VCCS 2010c). Since all pay raises in the VCCS are based on a percentage of a faculty member's current salary, any raise below the maximum is carried forward each successive year. In addition, the pay scale for hiring new faculty always increases at the same rate as the pay increase authorized by the state legislature for continuing faculty (VCCS, 2009). If a faculty member receives less than the legislated raise percentage, this may result in a situation where his pay level will lag behind his current colleagues' pay for the remainder of his career, and it may even put his salary below that of incoming faculty members who have less experience. As a result of these economic factors, the faculty evaluation process has ceased to serve one of the primary functions for which it was created – to provide formative feedback to faculty members in order to facilitate the process of continuous improvement.

According to C. D. Lee, Associate Vice Chancellor of Human Services for the VCCS (personal communication, March 20, 2008), between 2005 and 2007, 83% of the VCCS faculty members received “excellent” evaluation ratings and another 15% received “very good” ratings. Such high rankings could lead one to the conclusion that supervisors feel that there is not much need for improvement by the vast majority of their faculty. Instead the process has become almost entirely a high stakes summative prerequisite for faculty promotion and/or pay increase – or to serve as the springboard for a faculty

grievance when promotion and/or pay increase is held back due to a less-than-excellent evaluation rating.

As reported by Summers (2007), a second major limitation in the VCCS Policy Manual section on faculty evaluation is there is no link to the preceding section of the VCCS Policy Manual on “faculty responsibilities” (VCCS 2010a). Faculty responsibilities (see Appendix E for Sections 3.5 and 3.6 of the Policy Manual) include: (a) maintaining a “full” teaching load of 12-15 credit hours per semester, (b) having 10 or more office hours per week, (c) participating in “additional activities” such as committee work, student activities, student advising, etc., and (d) participation in other college, community or professional activities. This last area of responsibility reads, in part:

This requires that faculty members maintain current competence in their disciplines or specializations and that they share their expertise, time, and talents with the larger college community. Performance in this category will be measured not only by membership or affiliation but also by the quality of the contributions made by faculty members toward these endeavors (pp. 29-30).

This section of the VCCS Policy Manual continues with a list of thirteen examples of activities that could be considered under this responsibility area. However, the very next section of the VCCS Policy Manual does not explicitly address these faculty responsibilities, nor what portion of a faculty member’s summary evaluation rating should be tied to these responsibilities. The actual wording of the “Teaching Effectiveness” policy (3.6.0) is: “Components of teaching effectiveness *may include* but are not limited to: (a) performance in the classroom, (b) continuous updating, improvement, and innovation in teaching materials, methods, and assignments, (c)

maintenance of regular office hours, and (d) advisement of students” (VCCS 2010a, p. 3-33). In effect, the *required* teaching responsibilities from section 3.5 were “down-graded” to *optional* components of teaching effectiveness in the faculty evaluation section of the *VCCS Policy Manual*. As Summers (2007) concluded, there needs to be a better linkage between sections 3.5 and 3.6 of the *VCCS Policy Manual*.

Variations in the VCCS evaluation system. There are almost as many different ways to conduct a faculty evaluation as there are institutions that employ faculty. While faculty evaluation plans may be reduced to a few, common elements (student, peer, self, and administrative), the manner in which some or all of these elements are combined to produce the final result is almost infinitely variable. Each college in the VCCS has its own, unique, faculty evaluation plan – each of which came from the same basic guidelines in the *VCCS Policy Manual*. These guidelines only specify that “Performance evaluations shall include a summary rating of Excellent, Very Good, Good, Fair or Unsatisfactory” (VCCS, 2010a, p. 3-34), and that there are a minimum of four criteria which must be included in the plan: (a) effectiveness in performing tasks appropriate to their position description; (b) effectiveness in maintaining positive professional relationships; (c) effectiveness in maintaining current competence in their field; and (d) adherence to the policies and regulations of the VCCS and their respective colleges (p. 3-34).

“Effectiveness” was left for each college to determine as part of the development and implementation of their plan. Faculty evaluation plans must be approved by the college president and a majority of the faculty members at each college. Once approved at the local level, the plan also must be approved by the VCCS Human Resources office

staff. Nowhere does the *VCCS Policy Manual* specify what components must be a part of the faculty evaluation plan.

With over thirty years of research and publication experience with faculty evaluation, Raoul Arreola has developed a systematic process that has been used by personnel at hundreds of colleges and universities to develop comprehensive faculty evaluation plans for their institutions. His book, *Developing a Comprehensive Faculty Evaluation System: A Guide to Designing, Building, and Operating Large-Scale Faculty Evaluation Systems* (2007) provides a step-by-step recipe that can easily be followed and replicated. One fact that is emphasized repeatedly throughout Arreola's work is that there is no "perfect" plan that can be universally implemented. Each department in an institution has its own specific characteristics and mission, and faculty members in each department have different roles. Therefore, when it comes to the development of a comprehensive faculty evaluation system, one size does NOT fit all (Arreola; Centra, 1993; Seldin, 1999), and the construction of such a system is a time-consuming and difficult process. However, it is well worth the investment of time, money and effort to develop a strong system that is as fair and objective as possible (Arreola; Centra; Seldin). Therefore, while the VCCS faculty evaluation system requires substantial improvement, the literature suggests that variation in VCCS evaluation systems is a good thing and local control should remain a college issue.

Influence of the Literature on this Study

Based upon information in the literature, it is evident that faculty evaluation serves several, related purposes and that there is no single faculty evaluation plan that is appropriate for *all* colleges (Arreola, 2007; Centra, 1993; Seldin, 1999). Therefore, this

study explored areas of VCCS faculty and administrator perceptions with respect to: (a) the purposes of the faculty evaluation process, (b) the most important elements of a comprehensive faculty evaluation plan, (c) the perceived strengths and limitations of the current VCCS faculty evaluation process, (d) what should be changed in the current VCCS faculty evaluation process, and (e) how perceptions differed with respect to the survey participants' demographic and background variables (such as gender, primary role, primary teaching/administrative area, and years of experience).

In addition to collecting, analyzing, and reporting the results of this survey, one of the main goals of this study was to provide data to the colleges for use in the development/revision of their comprehensive faculty evaluation systems. It will be up to the individual college staff members whether or not they choose to utilize this data as they evaluate their existing plans for revision. On a more broad scale, the results of this study were used to create a "template" faculty evaluation plan for the VCCS that is built upon two foundation components: (a) best practices in the literature, and (b) actual data from VCCS teaching faculty members and their administrative supervisors.

This study makes two contributions to the faculty evaluation research literature by (1) adding to the very small number of studies conducted specifically on community college faculty and administrators and (2) exploring faculty opinions based on demographic subgroups. Due to the emphasis placed on teaching as the primary function for community college faculty, patterns discovered in evaluation studies on four-year faculty members may not necessarily translate well to community college faculty. There is essentially no literature concerning faculty demographic differences based on teaching area, years of experience, or gender.

CHAPTER 3

METHODOLOGY

This chapter identifies the purpose of the study, the design of the study, the population from which the sample was drawn, how the sample was obtained, the research questions, and the survey instrument developed to address the research questions. Validity and reliability of the survey instrument are described as well as the data collection procedure. Data analysis was performed utilizing descriptive statistics, multivariate analysis of variance (MANOVA) followed by univariate analysis of variance (ANOVA), and post-hoc contrasts (Scheffe) where there were more than two levels of the independent variable. Content analysis was performed on the qualitative data.

Research Design and Questions

A non-experimental design employing a survey instrument was used to address the research questions. The survey instrument was constructed following a blueprint that aligned the scales and items with the research questions, and utilized strategies to elicit both quantitative and qualitative responses.

Variables. The independent variables in this study were demographic in nature and included the following categories: *primary role* (teaching faculty or administrator), *primary content area* (developmental, career and technical, or transfer), *years of full-time experience*, and *gender* (female or male). The dependent variables were: (a) the faculty perceptions of the scaled scores related to the *purpose* of faculty evaluation, (b) the *elements* that should be included in the faculty evaluation process, (c) the *strengths and limitations* of the faculty evaluation process, and (d) the *changes needed* in the process.

Research Questions

1. What do the VCCS faculty members and administrators perceive to be the primary purposes of faculty evaluation (i.e. formative vs. summative)?
2. What do VCCS faculty members and administrators perceive to be the most important elements of a comprehensive faculty evaluation plan?
3. Do perceptions of the most important components used in the evaluation process differ as a function of demographic and background variables (i.e. primary role, primary content area, years of experience, and/or gender)?
4. What do VCCS faculty members and administrators perceive to be the strengths and limitations of the current faculty evaluation process?
5. What do VCCS faculty members and administrators suggest should be changed in the current faculty evaluation process?

Participants

Survey Population

The survey population consisted of all full-time teaching faculty members employed by the VCCS and their administrative supervisors, usually the academic deans. There are twenty-three colleges in the VCCS that occupy a total of 40 campus and numerous additional off-campus centers. The total academic student population for academic year 2007-2008 included almost 250,000 individual students (not including the 190,000 workforce development participants), who equated to nearly 102,000 full-time equivalent students (FTES). See Table 2 for college and system demographics.

Table 2

VCCS College Demographics for 2007-2008.

College Name	Number of Campuses	Student Headcount ^a	Student FTES ^a	Number of Full-time Teaching Faculty ^b	Number of Academic Deans ^c
Blue Ridge	1	5,765	2,623	67	3
Central Virginia	1	7,095	2,466	61	3
Dabney S. Lancaster	1	1,955	768	22	3
Danville	1	6,491	2,508	63	4
Eastern Shore	1	1,215	542	16	1
Germanna	2	8,184	3,368	58	2
J. Sargeant Reynolds	3	18,685	6,882	134	4
John Tyler	2	11,575	4,484	77	4
Lord Fairfax	2	7,669	3,076	62	5
Mountain Empire	1	4,312	1,918	46	3
New River	1	7,240	2,851	57	2
Northern Virginia	6	64,454	27,725	551	14
Patrick Henry	1	4,203	1,844	49	3
Paul D. Camp	2	2,318	869	20	2
Piedmont Virginia	1	6,598	2,456	67	3
Rappahannock	2	4,439	1,598	33	2
Southside Virginia	2	7,980	3,188	84	3
Southwest Virginia	1	5,807	2,324	69	6
Thomas Nelson	2	13,932	5,540	108	4

Tidewater	4	38,701	17,190	356	11
Virginia Highlands	1	3,532	1,572	47	3
Virginia Western	1	12,788	4,303	88	5
Wytheville	1	4,357	1,793	38	5
Totals	40	249,295	101,889	2,173	95

^a Enrollment figures from VCCS *Annual Enrollments by College* Report (VCCS, 2008b). ^b Faculty employment figures from VCCS *Federal Reports (IPEDS)* (VCCS, 2008c). ^c Dean employment figures from VCCS Council of Deans and Directors membership list (personal communication).

The potential study population included all academic deans and teaching faculty members from all 23 colleges in the VCCS as of spring semester 2009. There were approximately 95 academic deans and 2200 full-time teaching faculty members in the VCCS at that time (VCCS, 2008c). The actual study population did not include administrators or full-time teaching faculty members from some of the colleges since two college presidents did not agree to have their staff members contacted for participation in the study.

All of the academic deans were contacted directly via e-mail for participation in the study. The e-mail request for participation sent to the deans (see Appendix B) asked the deans to forward the participation request on to all teaching faculty members under their individual supervision. Although the total number of academic deans in the VCCS is relatively small, their response rate was very high. The teaching faculty population tends to be less responsive to invitations for survey participation but, due to the large size of this population, the sample was large. The sample is described in the next chapter.

Measure

Survey Instrument

An on-line survey instrument (Appendix C) was created and posted on a secure web server. The survey collected demographic information on the respondents appropriate to their primary role either as administrators or teaching faculty members, primary instructional area, years of experience in their current role, and gender. The survey was constructed to branch depending on whether the respondent indicated that s/he was a teaching faculty member or an administrator. Demographic questions varied based on the respondent's role. Following the demographic questions, all respondents were asked to respond to identical faculty evaluation close-ended and open-ended questions.

The survey instrument design. The survey instrument contained 68 questions, divided into six sections with scaled response options and one additional section with four open-ended questions. The instrument blueprint appears in Table 3. Other than the demographic questions, all of the other close-ended items had 4-point Likert type rating scale options. According to Berk (2006) an even-numbered scale, such as the four-point scale used for most sections of the survey instrument, removed the "neutral" middle position response that essentially would have allowed respondents to provide no useful information when asked to render an informed opinion. The four-point rating scales used in the survey instrument assessed (1) level of agreement: strongly disagree, disagree, agree and strongly agree, and (2) level of importance: not important, somewhat important, important, and very important. The first response scale was intended to quantify the respondents' level of agreement with the various purposes and uses of the faculty evaluation results, while the second scale was intended to quantify the level of

importance that the respondents placed on inclusion of each of the numerous components in a faculty evaluation plan.

Table 3

Blueprint of Faculty Evaluation Survey Instrument.

Content Area	Number of Items
Demographic Information	5
Purpose and Use of an Ideal Faculty Evaluation Plan	Level of agreement
Purpose and use of the faculty evaluation plan	7
Formative Evaluation of Overall Faculty Performance	Relative importance
Teaching	1
Service to students/college/community	1
Scholarly/creative activities	1
Professional recognition/accomplishment	1
Personal attributes	1
Other	1
Summative Evaluation of Overall Faculty Performance	Relative importance
Teaching	1
Service to students/college/community	1
Scholarly/creative activities	1
Professional recognition/accomplishment	1
Personal attributes	1
Other	1
Evaluation of Teaching Performance	Relative importance
Classroom performance	5

Preparation for class/laboratory/clinical/etc.	6
Student performance & evaluation of faculty	4
Self evaluation	3
Additional/external evaluation	7
<hr/>	
Evaluation of Service Performance	Relative importance
Service to the students	6
Service to the college	4
Service to the community	5
<hr/>	
Summary Comments on the Faculty Evaluation Process	Open-ended questions
Summary comments on the faculty evaluation process	4
<hr/>	
Total Number of Questions	68
<hr/>	

Validity. The literature review provided the foundation for the survey instrument with selected questions modified from Peter Seldin's 1983 survey instrument, Raoul Arreola's 2007 "partial list of possible faculty roles," and additional questions developed by the investigator. Therefore, the blueprint based on the literature enhanced content validity. Three experts in the field of performance evaluation and survey methodology were asked to review the survey to further enhance validity, and revisions were made based on the feedback of the expert reviewers. Factor analysis was conducted on the survey results to estimate factorial validity.

Reliability. To estimate reliability of the survey instrument, Cronbach's alpha analysis was conducted on each scale. Two individuals coded responses for the open-ended questions and inter-rater reliability was calculated using Cohen's Kappa for level

of agreement on the coded responses. A minimum of 80% agreement between the raters for the coding of responses provided additional evidence of data reliability.

Procedure

Administration

The process for collecting data began with a letter sent by the investigator to the VCCS presidents seeking permission to contact their academic deans and faculty members with a request to participate in the study (see Appendix A). Upon receiving permission from the college president, an invitation to participate in the study was e-mailed to their academic deans for their use and for forwarding on to their full-time teaching faculty members (Appendix B). The invitation contained a link to the survey instrument (Appendix C) as well as a brief description of the study and a notification that all responses would remain anonymous with only aggregated data released to the colleges or included in publications.

To encourage individuals to participate in the survey, they were assured of anonymity as no names or e-mail address of the respondents were collected or tied to any survey responses. While there were identifiers related to the participants' demographic information (college, gender, primary role, years of experience, and primary content area) no attempt was made to relate any of these responses to individual participants and only grouped data was reported. Therefore, no individuals were identified in any published documents or communications between the author and any other individuals. The author also offered participants an opportunity to enter into a lottery drawing for 12, \$25 gift cards through a separate on-line site that was available through a link at the end of the survey instrument (Appendix D).

As yet another incentive for participation, participants were informed they would have access to anonymous data for strategic planning purposes or simply for their information. Grouped data will be sent to the participating college presidents after the completion of the study. Only system-wide aggregate data will be provided to the presidents at the participating colleges and the aggregated data will be posted on a separate web site for open access by deans, faculty members or any other interested individuals. College personnel may choose to use grouped response data to help shape the content and emphases of their revised faculty evaluation plans, but there will not be any VCCS requirement to do so.

Follow-up

The investigator periodically monitored the number of survey participants from each of the participating colleges. As the deadline for participation approached the final five days a targeted second appeal was sent to each of the deans at the colleges for which the faculty response was less than 20% of the prospective participants. This second request led to additional participation at each of these colleges.

Data Analysis

Primary Purpose of Faculty Evaluation

The data analysis was both descriptive and comparative. The survey was divided into four major sections. The first major section of the survey instrument was used to collect demographic data, which served as the basis for comparing different demographic subsets of the respondents for quantitative analysis.

The second section was composed of three subsections (“Purpose and Use of an Ideal Faculty Evaluation Plan,” “Formative Evaluation of Overall Faculty Performance,”

and “Summative Evaluation of Overall Faculty Performance”) that were intended to provide data for comparisons between faculty and administrator responses related to the purpose(s) for conducting an annual faculty evaluation. Survey items prompted respondents to rank their level of agreement as to whether or not faculty evaluation should be used for formative and/or summative purposes, and to identify the main components that should be part of this evaluation. This section of the survey instrument was analyzed using multivariate analysis of variance (MANOVA) to look for differences between faculty and administrator responses for each section and univariate analysis of variance (ANOVA) calculations were performed to compare faculty and administrator response means when there were significant MANOVA results. Descriptive statistics were also calculated for each individual item/dependent variable for which there were significant ANOVA results.

The third major section of the survey instrument contained two subsections (“Evaluation of Teaching Performance” and “Evaluation of Service Performance”) that went into more detailed analysis of some of the individual items (dependent variables) that contribute to the overall formative and summative evaluation components. Factor analysis was utilized to compare demographic group responses to these survey items, resulting in nine scales that were further analyzed using multivariate analysis of variance (MANOVA) calculations to identify significant differences between the various demographic subsets of the study sample respondents (i.e., the independent variables). When significant demographic differences were identified through the MANOVA calculations, additional univariate analysis of variance (ANOVA) calculations were

performed to identify the specific differences between demographic group mean responses on the dependent variables (i.e. the nine scales).

The fourth, and final, section of the survey instrument was for open-ended “Summary Comments on the Faculty Evaluation Process.” Qualitative content analysis was used to examine responses to this portion of the survey instrument.

CHAPTER 4

RESULTS

Following a brief description of the study population, quantitative results are presented for the comparison of faculty and administrator perceptions of the purpose and use of the faculty evaluation process for formative and summative purposes. Factor analysis of teaching and service performance aspects of the faculty evaluation process is described as a basis for comparing faculty and administrator responses as well as for comparing faculty responses based on their primary teaching area, years of experience, and gender. Qualitative content analysis is then presented regarding strengths, limitations, and suggested changes for existing faculty evaluation plans.

Table 4 presents the demographic data for the sample, which consisted of 67 administrators and 404 full-time teaching faculty members for a total of 471 respondents. Using college employee data from the 2007-2008 academic year, the most current data available when the survey was conducted, the potential survey population consisted of 1756 full-time teaching faculty members and 81 academic deans. Therefore, the response rate represents 23.0% of the eligible teaching faculty members and 82.7% of the eligible academic deans from the survey population. Of the 404 teaching faculty respondents, 170 taught primarily transfer courses, 175 taught career and technical courses, 29 taught developmental courses, and 10 taught “other” courses. The largest number of faculty respondents taught for less than seven years (146) or more than 15 years (145), with the minority (112) having 7 – 15 years of experience. More female faculty members (229) than male faculty members (172) participated in the study.

Table 4

VCCS College Demographics for 2007-2008 and Number of Survey Respondents.

College Name	Number of Full-time Teaching Faculty ^a	Number of Academic Deans ^b	Number of Full-time Teaching Faculty Responding to the Survey	Number of Academic Deans Responding to the Survey
Blue Ridge	67	3	20	3
Dabney S. Lancaster	22	3	10	2
Danville	63	4	33	4
Eastern Shore	16	1	5	1
Germanna	58	2	25	2
J. Sargeant Reynolds	134	4	43	3
John Tyler	77	4	25	4
Lord Fairfax	62	5	25	5
Mountain Empire	46	3	9	3
New River	57	2	15	2
Northern Virginia	551	14	35	4
Patrick Henry	49	3	12	3
Paul D. Camp	20	2	11	1
Piedmont Virginia	67	3	9	3
Rappahannock	33	2	17	2
Southside Virginia	84	3	22	3
Southwest Virginia	69	6	15	6

Thomas Nelson	108	4	21	4
Virginia Highlands	47	3	10	2
Virginia Western	88	5	31	5
Wytheville	38	5	11	5
Totals	1756	81	404	67

^aFaculty employment figures from VCCS *Federal Reports (IPEDS)* (VCCS, 2008c). ^bDean employment figures from VCCS Council of Deans and Directors membership list (personal communication).

Quantitative Results

Research question 1. What do the VCCS faculty members and administrators perceive to be the primary purposes of faculty evaluation (i.e. formative vs. summative)?

For the first three objective (non-demographic) parts of the survey instrument “Purpose and Use of an Ideal Faculty Evaluation Plan,” “Formative Evaluation of Overall Faculty Performance,” and “Summative Evaluation of Overall Faculty Performance,” there were not enough related items in these sections of the survey instrument to provide any meaningful factor analysis results. Therefore, descriptive statistics were employed to reveal the purposes and components of the faculty evaluation process. Analysis of variance calculations were performed to identify significant demographic group differences and descriptive statistics were performed on close-ended items on the survey instrument. Descriptive statistics included the means and standard deviations for all respondents as well as the faculty and administrator response means for the survey items for which there were significant findings.

Purpose and use of a faculty evaluation plan. The four-point “agreement” rating scale used in the “Purpose and Use of an Ideal Faculty Evaluation Plan” portion of the

survey instrument ranged from “strongly disagree” (1) to “strongly agree” (4). This response scale was intended to quantify the respondents’ level of agreement with the various purposes and uses of the faculty evaluation results. Response means were calculated for each item on the survey.

The descriptive statistics for all responses are presented in Table 5 and have been ranked in descending order based on the overall response means. Response means indicate an overall level of agreement to the specified purpose or use of faculty evaluations, with higher means denoting stronger agreement. As detailed in Table 5, every item had an overall mean greater than 2.50 on the agreement scale.

Table 5

Purpose and Use of an Ideal Faculty Evaluation Plan: Overall Means

Item	N ^a	Overall Mean	Overall Std Dev
Both formative and summative process	450	3.16	.822
Evaluators should have discretion	450	3.09	.590
Primarily a formative process	450	3.08	.802
Tied to professional development	450	2.99	.690
2 or more excellent for promotion	450	2.95	.806
Merit pay tied to summary ratings	450	2.71	.795
Primarily a summative process	450	2.54	.836

^a386 faculty + 64 administrators responded to all items in this portion of the survey instrument

Collectively, the survey respondents had the strongest rating (3.16) for using the faculty evaluation plan for both formative and summative purposes, with four other items

receiving slightly lower ratings. Formative plans are designed to enhance continuous improvement in performance of teaching faculty members, while summative plans are designed to provide an overall performance rating to be used by administrators as the basis for making personnel decisions such as determining contract eligibility, promotion, and merit pay status.

Group means were compared through multivariate analysis of variance (MANOVA) to see if there were significant differences between the faculty and administrator responses to this portion of the survey instrument. The results revealed a statistically significant main effect for the purpose and uses of faculty evaluation, $F(1, 449) = 7.91, p < .001$, Wilks' $\lambda = .889$, with respect to faculty and administrator responses on this part of the survey instrument. Univariate analysis of variance (ANOVA) calculations were then conducted to compare the means between the faculty and administrator responses. As shown in Table 6, there were five significant, $p < .05$, results identified for the purpose and use of faculty evaluation. The two items with the lowest overall means were the two concepts for which there was no significant difference between administrator and faculty responses: (a) merit pay tied to summary ratings (2.71) and (b) primarily a summative process (2.54).

Table 6

Purpose and Use of an Ideal Faculty Evaluation Plan: Faculty vs. Administrator, ANOVA

Survey Item	Between-Groups F	Significance
Tied to professional development	34.182*	.000
Evaluators should have discretion	10.529*	.001

Both formative and summative process	8.941*	.003
2 or more excellent ratings for promotion	7.528*	.006
Primarily a formative process	7.233*	.007
Merit pay tied to summary ratings	1.689	.194
Primarily a summative process	.308	.579

*Significant at $p < .05$

Faculty and administrator response means for the significant differences are compared in Table 7. As shown in Table 7, administrators rated all items higher than faculty members with professional development having the greatest difference (0.53).

Table 7

Purpose and Use of an Ideal Faculty Evaluation Plan: Mean Comparisons

Item	Faculty Mean	Admin Mean	Mean Difference (Admin – Faculty)
Tied to professional development	2.91	3.44	0.53
Both formative and summative process	3.11	3.44	0.33
Primarily a formative process	3.04	3.33	0.29
2 or more excellent for promotion	2.91	3.20	0.29
Evaluators should have discretion	3.06	3.31	0.25

Research question 2. What do VCCS faculty members and administrators perceive to be the most important elements of a comprehensive faculty evaluation plan?

The four-point rating scale used in the “Formative Evaluation of Overall Faculty Performance” and “Summative Evaluation of Overall Faculty Performance” portions of

the survey instrument ranged from “not important” (1) to “very important” (4). This scale was intended to quantify the level of importance that the respondents placed on using each of the numerous components that could be included in a faculty evaluation plan.

Formative evaluation of overall faculty performance. As shown in Table 8, teaching received the highest overall rating ($M = 3.90$) by all respondents, with all other items receiving substantially lower response means. Professional recognition ($M = 2.37$) and scholarly or creative activities ($M = 2.15$) received the lowest overall response means.

Table 8

Overall Formative Faculty Performance: Overall Means and Standard Deviations

Item	N^a	Overall Mean	Overall Std Dev
Teaching	467	3.90	.362
Personal attributes	467	3.03	.814
Service	467	2.82	.729
Professional recognition	467	2.37	.788
Scholarly or creative activity	467	2.15	.794

^a400 faculty + 67 administrators responded to all items in this portion of the survey instrument

Multivariate analysis of variance (MANOVA) calculations were performed to test if there were significant differences between the faculty and administrator responses on this portion of the survey instrument. The results revealed a statistically significant main effect for the formative applications of faculty evaluation, $F(1, 466) = 4.41, p = .001$, Wilks' $\lambda = .954$, with respect to faculty and administrator responses on this part of the survey instrument.

Univariate analysis of variance (ANOVA) calculations were then conducted to compare the means between the faculty and administrator responses. There were two significant differences, related to the use of service and scholarly or creative activities for formative evaluation, as shown in Table 9.

Table 9

Overall Formative Faculty Performance: Faculty vs. Administrator ANOVA

Survey Item	Between-Groups <i>F</i>	Significance
Service	13.145*	.000
Scholarly or creative activity	12.450*	.000
Personal attributes	3.111	.078
Professional recognition	2.840	.093
Teaching	.002	.967

*Significant at $p < .05$

Table 10 shows that for the two items where the administrators and faculty members differed in their response means, the administrators had the higher mean responses. There was no difference in the rating of teaching as the most important item for inclusion in a formative faculty evaluation plan. Faculty and administrators had the same mean response of 3.90.

Table 10

Overall Formative Faculty Performance: Mean Comparisons

Item	Faculty Mean	Admin Mean	Mean Difference (Admin – Faculty)
Scholarly or creative activity	2.10	2.46	0.36
Service	2.78	3.12	0.34

Summative evaluation of overall faculty performance. Regardless of the purpose for utilizing the results of the plan (formative or summative) the highest overall ratings were for “teaching.” Teaching received the highest mean for summative evaluation ($M = 3.80$) with professional recognition ($M = 2.43$) and scholarly or creative activities ($M = 2.22$) again having the lowest means as shown in Table 11. These results also support the statements made earlier in this dissertation concerning the overwhelming emphasis on teaching for community college faculty members versus scholarly or creative activities that may be much more important for faculty members at senior institutions. As with the responses for formative evaluation (Table 8), faculty and administrators also ranked these five items in the same order of importance for summative evaluation.

Table 11

Overall Summative Faculty Performance: Overall Means and Standard Deviations

Item	N^a	Overall Mean	Overall Std Dev
Teaching	459	3.80	.444
Personal attributes	459	3.01	.849
Service	459	2.84	.726

Professional recognition	459	2.43	.802
Scholarly or creative activity	459	2.22	.812

^a392 faculty + 67 administrators responded to all items in this portion of the survey instrument

Multivariate analysis of variance (MANOVA) calculations were performed to see if there were significant differences between the faculty and administrator responses for this portion of the survey instrument. The results revealed a statistically significant main effect for the summative applications of faculty evaluation, $F(1, 458) = 6.94, p < .001$, Wilks' $\lambda = .929$, with respect to faculty and administrator responses on this part of the survey instrument.

Univariate analysis of variance (ANOVA) calculations were then conducted to compare the means between the faculty and administrator responses. For the univariate analysis of the formative evaluation data, the two significant differences were for the uses of service and scholarly or creative activities in the summative evaluation, as shown in Table 12.

Table 12

Overall Summative Faculty Performance: Faculty vs. Administrator ANOVA

Survey Item	Between-Groups F	Significance
Service	28.886*	.000
Scholarly or creative activity	5.549*	.019
Teaching	3.354	.068
Personal attributes	.999	.318
Professional recognition	.312	.577

*Significant at $p < .05$

Here, as for the formative evaluation data, the responses show that administrators gave more importance to the two significant items than did the faculty members, although both groups gave these variables similar levels of importance, as shown in Table 13.

Table 13

Overall Summative Faculty Performance: Mean Comparisons

Item	Faculty Mean	Admin Mean	Mean Difference (Admin – Faculty)
Service	2.77	3.27	0.50
Scholarly or creative activity	2.18	2.43	0.25

Evaluation of teaching performance activities. Since teaching was anticipated to be the most important component for inclusion in the faculty evaluation process, several additional items were included in the survey instrument to further explore which particular responsibilities related to teaching VCCS faculty and administrators identified as being the most important aspects for inclusion in the faculty evaluation plan. For the teaching performance activities detailed in Table 14, the three activities rated with the highest overall response means were (a) developing course materials ($M = 3.38$), (b) course syllabi and examinations ($M = 3.33$), and (c) delivering lectures ($M = 3.27$). Other performance measures dropped in importance compared with these top three items with enrollment in elective courses ($M = 1.95$) and operating a chat room or discussion board ($M = 1.92$) receiving the lowest response means.

Table 14

Teaching Performance Activities: Overall Means and Standard Deviations

Item	N ^a	Overall Mean	Overall Std Dev
Developing course materials	440	3.38	.697
Course syllabi and examinations	440	3.33	.691
Delivering lectures	440	3.27	.762
Grading examinations	440	3.12	.806
Supervising laboratory sessions	440	3.04	.845
Developing written examinations	440	3.00	.796
Course load	440	2.98	.872
Facilitating small group experiential learning	440	2.50	.893
Creating an on-line course	440	2.19	.917
Enrollment in elective courses	440	1.95	.875
Operating a chat room or discussion board	440	1.92	.860

^a379 faculty + 61 administrators responded to all items in this portion of the survey instrument

Multivariate analysis of variance (MANOVA) calculations were performed to see if there were significant differences between the faculty and administrator responses for the items in this portion of the survey instrument. The results revealed a statistically significant main effect for the teaching performance activities that are often included in faculty evaluation, $F(1, 439) = 6.03, p < .001$, Wilks' $\lambda = .866$, with respect to faculty and administrator responses on this part of the survey instrument.

Univariate analysis of variance (ANOVA) calculations were then conducted to compare the means between the faculty and administrator responses. For the univariate analysis of the teaching performance activities data, there were two significant classroom performance differences (operating a chat room or discussion board and facilitating small

group experiential learning groups) and four significant differences for activities related to preparation for class/laboratory/clinical/etc. as shown in Table 15.

Table 15

Teaching Performance Activities: Faculty vs. Administrator, ANOVA

Survey Item	Between-Groups <i>F</i>	Significance
<i>Classroom performance</i>		
Operating a chat room or discussion board	35.115*	.000
Facilitating small group experiential learning events	15.696*	.000
Enrollment in elective courses	1.671	.197
Supervising laboratory sessions	1.504	.221
Delivering lectures	.993	.320
<i>Preparation for class/laboratory/clinical/etc.</i>		
Developing course materials	8.954*	.003
Creating an on-line course	8.747*	.003
Grading examinations	7.145*	.008
Course load	7.004*	.008
Course syllabi and examinations	2.971	.085
Developing written examinations	.270	.604

*Significant at $p < .05$

Table 16 shows the faculty and administrator mean comparisons. Operating a chat room or discussion board showed the greatest difference between faculty and administrator response means (0.68), followed by facilitating small group experiential

learning activities (0.48) and creating an on-line course (0.37). All three of these activities could be considered “non-traditional” teaching activities and all three received higher response means from administrators. Course load was the only teaching performance activity where the faculty response mean was greater than the administrator mean.

Table 16

Teaching Performance Activities: Mean Comparisons

Item	Faculty Mean	Admin Mean	Mean Difference (Admin – Faculty)
Operating a chat room or discussion board	1.83	2.51	0.68
Facilitating small group experiential learning	2.44	2.92	0.48
Creating an on-line course	2.14	2.51	0.37
Course load	3.02	2.70	-0.32
Grading examinations	3.08	3.38	0.30
Developing course materials	3.34	3.62	0.28

In addition to teaching performance activities, respondents were asked to rate the importance of including various contributing sources when constructing the evaluation rating. As shown in Table 17, only one contributing source elicited an overall importance mean in excess of 3.0, the faculty member’s direct supervisor ($M = 3.04$). In descending order of overall means, the next three highest rated sources were (a) dean evaluation ($M = 2.77$), (b) systematic student evaluation ($M = 2.69$), and (c) self evaluation checklist or rating scale ($M = 2.61$). Response means dropped for the remaining contributing sources,

with alumni evaluation ($M = 1.69$) and committee evaluation ($M = 1.68$) having the lowest means.

Table 17

Contributing Sources for Teaching: Overall Means and Standard Deviations

Item	N^a	Overall Mean	Overall Std Dev
Direct supervisor's evaluation	442	3.04	.694
Dean evaluation	442	2.77	.827
Systematic student evaluations	442	2.69	.822
Self evaluation checklist or rating scale	442	2.61	.784
Self evaluative narrative report	442	2.56	.804
Student examination performance	442	2.51	.768
Peer evaluation	442	2.33	.846
Compilation of detailed portfolio	442	2.31	.948
Informal student opinions	442	2.29	.775
Grade distributions	442	2.05	.751
VP/Provost evaluation	442	1.93	.805
Alumni evaluation	442	1.69	.787
Committee evaluation	442	1.68	.754

^a383 faculty + 59 administrators responded to all items in this portion of the survey instrument

Multivariate analysis of variance (MANOVA) calculations were performed to see if there were significant differences between the faculty and administrator responses for the items in this portion of the survey instrument. The results revealed a statistically significant main effect for the contributing sources that are often included in faculty

evaluation, $F(1, 441) = 3.70$, $p < .001$, Wilks' $\lambda = .899$, with respect to faculty and administrator responses on this part of the survey instrument.

Univariate analysis of variance (ANOVA) calculations were then conducted to compare the means between the faculty and administrator responses. For the univariate analysis of the teaching performance activities data, there were nine significant differences between faculty and administrator responses as shown in Table 18. All four items in the student performance and evaluation of faculty section were significantly different, one of three self evaluation items (compilation of a detailed portfolio) was significantly different, and four of the six additional/external sources of evaluation were rated as significantly different by the faculty and administrators.

Table 18

Contributing Sources: Faculty vs. Administrator, ANOVA

Survey Item	Between-Groups F	Significance
<i>Student performance and evaluation of faculty</i>		
Systematic student evaluations	24.135*	.000
Student examination performance	6.415*	.012
Grade distributions	4.869*	.028
Informal student opinions	3.904*	.049
<i>Self evaluation</i>		
Compilation of a detailed portfolio	4.841*	.028
Self evaluative narrative report	2.333	.127
Self evaluation checklist or rating scale	.466	.495

<i>Additional/external evaluation</i>		
Dean evaluation	24.407*	.000
Direct supervisor's evaluation	20.010*	.000
VP/Provost evaluation	10.293*	.001
Committee evaluation	7.956*	.005
Peer evaluation	.829	.363
Alumni evaluation	.206	.650

*Significant at $p < .05$

Table 19 shows the comparison of the faculty and administrator response means for the significant findings. Administrators rated every one of these contributing sources higher than faculty members with dean evaluation and systematic student evaluations sharing the greatest mean difference (0.55) followed by the direct supervisor's evaluation (0.43).

Table 19

Contributing Sources for Teaching: Mean Comparisons

Item	Faculty Mean	Admin Mean	Mean Difference (Admin – Faculty)
Dean evaluation	2.70	3.25	0.55
Systematic student evaluations	2.62	3.17	0.55
Direct supervisor's evaluation	2.98	3.41	0.43
VP/Provost evaluation	1.88	2.24	0.36
Compilation of detailed portfolio	2.27	2.56	0.29
Committee evaluation	1.64	1.93	0.29

Student examination performance	2.48	2.75	0.27
Grade distributions	2.02	2.25	0.23
Informal student opinions	2.26	2.47	0.21

In summary, significant multivariate analysis of variance (MANOVA) results were obtained for each of the above content areas on the survey instrument with respect to faculty and administrator differences. Wilks' λ results for each MANOVA calculation in this section are presented in Table 20. Each MANOVA calculation resulted in a significant Wilks' λ , $p < .05$.

Table 20

Summary Wilks' λ Results for all Faculty vs. Administrator MANOVA Calculations

Survey Instrument Content Area	λ	F	Significance
Purpose and use of an ideal faculty evaluation plan	.889	7.91*	.000
Formative evaluation of overall faculty performance	.954	4.41*	.001
Summative evaluation of overall faculty performance	.929	6.94*	.000
Evaluation of teaching performance activities	.866	6.03*	.000
Contributing sources for the evaluation of teaching performance	.899	3.70*	.000

* $p < .05$.

Factor Analysis

Factor analysis is a statistical technique that is useful when looking for underlying relationships between large numbers of variables. This technique allows the researcher to reduce a large number of variables into a smaller number of variables, or “factors,” that are representative of some unobserved connection between the original variables. While the first three objective (non-demographic) portions of the survey instrument did not contain a large enough number of related items for factor analysis to provide any meaningful results, factor analysis was performed on the 24 objective items included in Tables 14 and 17 from the “Evaluation of Teaching Performance” portion of the survey instrument to develop scales based on items related to perceptions of VCCS teaching faculty and administrators on the importance of including various activities/responsibilities and contributing data sources in the faculty evaluation process. Factor analysis was performed to confirm the construct validity of the survey instrument scales and to construct related-item scales for further analysis. All faculty and administrator responses were utilized for this portion of the data analysis.

Evaluation of Teaching Performance

The first factor analysis was conducted on the 24 items comprising the Evaluation of Teaching Performance section of the questionnaire. Initially, principle components extraction was performed with oblimin rotation to determine the number of components. The factor loading criterion for discarding items from further analysis was set at 0.30 for correlations between items and components. For the principal components analysis, no items had correlations less than 0.30. As a result of the factor analysis, eight components emerged from the activities contained in the “Evaluation of Teaching Performance”

portion of the survey instrument with eigenvalues greater than one. These eight components accounted for approximately 67% of the total variance.

Further analysis of the items that comprised each of these eight components using oblimin rotation and Kaiser normalization resulted in the reduction of the “Evaluation of Teaching Performance” components to six components, due to two situations: either (a) a specific items factored into two of the components and was removed from the component where it was a minor factor and/or (b) there were less than three items in the component. This reduction of components led to the removal of three items (enrollment in elective courses, delivering lectures, and supervising laboratory sessions) from the developing scales. As shown in Table 21, the resultant six factors accounted for approximately 57% of the total variance.

Table 21

Pattern Matrix Principal Component Analysis for “Evaluation of Teaching Performance” Items

Survey Item	<u>Factor</u>					
	1	2	3	4	5	6
<i>Student Performance</i>						
Informal student opinions	.779					
Systematic student evaluations	.736					
Student examination performance	.671					
Grade distributions	.461					
<i>Preparation for Class</i>						
Grading examinations		.789				
Developing written examinations		.761				

Course syllabi and examinations	.730					
Developing course materials	.620					
Course load	.535					
<i>Self Evaluation</i>						
Self-evaluation narrative report	.852					
Self-evaluation checklist or rating scale	.836					
Compilation of detailed portfolio	.515					
<i>Administrator Evaluation</i>						
Dean evaluation	.886					
Direct supervisor's evaluation	.824					
VP/Provost evaluation	.557					
<i>Alternative Delivery Methods of Instruction</i>						
Operating a chat room or discussion board	.840					
Creating an on-line course	.753					
Facilitating experiential learning events	.654					
<i>External Evaluation</i>						
Committee evaluation	.820					
Alumni evaluation	.749					
Peer evaluation	.491					
<i>Variance (%)</i>	24.2	8.8	7.3	6.5	5.7	4.7

The first factor consisted of a four-item “student performance” scale that focused on student evaluation of teaching performance and student performance in class. The second factor consisted of a five-item “preparation for class” scale that focused on faculty work outside of class time. The third factor consisted of a three-item “self evaluation”

scale that focused on the self-documentation of faculty accomplishments. The fourth factor consisted of a three-item “administrator evaluation” scale that focused on evaluations of faculty performance by direct-line supervisors. The fifth factor consisted of a three-item “alternative delivery methods of instruction” scale that focused on “non-traditional” methods of teaching. The sixth factor consisted of a three-item “external evaluation” scale focused on faculty evaluation by others less directly related to the faculty member’s performance activities. Cronbach’s alpha coefficients were computed to estimate the inter-item reliabilities within these scales, and the results supported the moderate reliabilities of the scales: student performance scale = .72, preparation for class scale = .79, self-evaluation scale = .67, administrator evaluation scale = .77, alternative delivery methods scale = .70, and external evaluation scale = .62. These scale scores were established using all respondents, then used as the basis for examining differences by demographic groups.

Evaluation of Service Performance

The second factor analysis was performed on the 14 objective items from the “Evaluation of Service Performance” portion of the survey instrument to develop scales based on items related to perceptions of VCCS teaching faculty and administrators on the service components used in the faculty evaluation process. Initially, principle components extraction was performed with oblimin rotation to determine the number of components. The criterion for discarding items from further analysis was set at 0.30 for correlations between items and components. For the principal components analysis, no items had correlations less than 0.30. As a result of the factor analysis, three components emerged from the items in the “Evaluation of Service Performance” portion of the survey

instrument with eigenvalues greater than one. These three components accounted for approximately 62% of the total variance.

Further analysis of the items that composed each of these three components using oblimin rotation and Kaiser normalization resulted in the retention of the three “Evaluation of Service Performance” components. This process led to the removal of one item (academic advising) from the developed scales since it loaded almost equally into two different factors. There was also a resultant switching of two items from one of the survey scales into a different scale. As shown in Table 22, “maintaining regular office hours” moved from the “service to students” scale on the survey instrument to the “service to the college” scale. Similarly, “service as a student recruiter” moved from the “service to the college” scale on the survey instrument to the “service to students” scale.

Table 22

Pattern Matrix Principal Component Analysis for “Evaluation of Service Performance” Items

Survey Item	<u>Factor</u>		
	1	2	3
<i>Service to the College</i>			
Service on department committees	.923		
Service on college committees	.879		
Maintaining regular office hours	.581		
Departmental administrative duties	.522		
<i>Service to the Community</i>			
Provision of discipline-related expertise to community groups		.836	
Service on local community boards		.822	

Participation in local activities/events	.806
Service as a judge for local civic or school events	.795
<i>Service to Students</i>	
Non-academic advising	.837
Advisor to student organizations	.710
Willingness to teach “undesirable” courses	.605
Service as student recruiter	.549
Participation in campus programs	.454
<i>Variance (%)</i>	40.1 12.4 9.6

The resultant three factors accounted for approximately 62% of the total variance. The first factor consisted of a four-item “service to the college” scale that focused on college-related duties separate from instruction. The second factor consisted of a four-item “service to the community” scale that focused on faculty work outside of the college and in the community. The third factor consisted of a five-item “service to students” scale that focused on interactions with students outside of the classroom setting. Cronbach’s alpha coefficients were computed to estimate the inter-item reliabilities within these scales, and the results supported the reliability of the scales: service to the college scale = .78, service to the community scale = .88, and service to students scale = .77.

Research question 3. Do perceptions of the most important components used in the evaluation process differ as a function of demographic and background variables (i.e. primary role, primary content area, years of experience, and/or gender)?

As described earlier in this chapter, factor analysis was performed to confirm the construct validity of the survey instrument scales and to construct related-item scales for further analysis. As a result of the factor analysis, the resultant six scales from the “Evaluation of Teaching Performance” portion of the survey instrument were labeled as follows: (a) student performance, (b) preparation for class, (c) self evaluation, (d) administrator evaluation, (e) alternative delivery methods of instruction, and (f) external evaluation. The three resultant scales from “Evaluation of Service Performance” of the survey instrument were labeled as follows: (g) service to the college, (h) service to the community, and (i) service to students.

To address this research question regarding differences between various demographic subsets of the sample population, descriptive statistics and between-subjects multivariate analysis of variance (MANOVA) calculations were performed. When the MANOVA calculations for specific demographic comparisons yielded significant results, univariate analysis of variance (ANOVA) calculations were performed to determine which specific dependent variables contributed to these demographic differences.

Primary Role. Primary role (i.e. faculty or administrator) was the subject of the first demographic comparison. As shown in Table 23, the two areas that received the highest overall response means were preparation for class ($M = 3.16$) and service to the college ($M = 2.95$). The response means for the remaining items decreased rapidly with service to the community ($M = 2.08$) and external evaluators ($M = 1.90$) receiving the lowest overall ratings.

Table 23

Primary Area of Responsibility: Overall Means and Standard Deviations

Item	N ^a	Overall Mean	Overall Std Dev
Prep for class	469	3.16	.574
Service to the college	469	2.95	.612
Administrator evaluation	469	2.59	.652
Self evaluation	469	2.50	.664
Service to students	469	2.47	.634
Student performance	469	2.39	.577
Alternative delivery formats	469	2.22	.718
Service to the community	469	2.08	.681
External evaluators	469	1.90	.606

^a402 faculty + 67 administrators responded to all items in this portion of the survey instrument

When compared through MANOVA calculations, there were significant differences of perception between the faculty and administrator groups, $F(1, 468) = 8.78$, $p < .001$, Wilks' $\lambda = .853$. Univariate analysis of variance (ANOVA) calculations were then performed as a follow-up to the significant findings from the MANOVA analysis on the nine scales for faculty performance to test for differences between the dependent variables (e.g. service to the college) with respect to each independent variable (e.g. primary role), and to determine which demographic group had the higher mean response for each of these nine scales. No post hoc analyses were performed for primary role differences since there were only two groups to compare (faculty and administrators). The results of the univariate analysis F-test for between-group mean comparisons for

significant differences are shown in Table 24. The Bonferroni error adjustment was made for multiple comparisons within the ANOVA calculations.

Table 24

Primary Area of Responsibility: Between-Subjects Effects, ANOVA

Scale	<i>F</i>	Significance**
Service to the college	36.934*	.000
Administrator evaluation	30.407*	.000
Alternative delivery formats	26.810*	.000
Service to students	18.222*	.000
Student performance	13.888*	.000
Service to the community	13.238*	.000
External evaluators	3.408	.066
Self evaluation	3.404	.066
Prep for class	1.545	.214

* $p < .05$. **Bonferroni adjustment for multiple comparisons.

As shown in Table 25, the response mean differences for the significant ANOVA findings were greatest for alternative delivery formats (.48), service to the college (.47), and administrator evaluation (.47). The other three areas of responsibility with significant ANOVA findings had lower mean response differences with student performance having the least difference (.28).

Table 25

Primary Area of Responsibility: Mean Comparisons

Item	Faculty Mean	Admin Mean	Mean Difference (Admin – Faculty)
Alternative delivery formats	2.15	2.63	0.48
Service to the college	2.88	3.35	0.47
Administrator evaluation	2.52	2.99	0.47
Service to students	2.42	2.77	0.35
Service to the community	2.03	2.35	0.32
Student performance	2.35	2.63	0.28

Faculty Demographic Analysis

Additional MANOVA calculations were performed to determine if significant differences exist between the VCCS faculty respondents from the different demographic groups (primary teaching area, years of full-time teaching experience, and gender) with respect to their perceptions of the faculty evaluation process relative to the nine faculty performance scales. Due to low sample sizes, the 29 faculty members who identified their primary responsibility area as “developmental” and the 10 faculty members who identified their primary responsibility area as “other” were removed from the analysis. Thus the comparisons were between those teaching primarily in either the transfer or the career and technical (CTE) content areas.

Primary teaching area. As shown in Table 26, preparation for class received the highest overall rating ($M = 3.14$) and service to the college ($M = 2.87$) was the only other item with an overall mean greater than 2.50. External evaluators ($M = 1.88$) was the only item to receive a mean response below 2.00.

Table 26

Primary Teaching Area: Overall Means and Standard Deviations

Item	N ^a	Overall Mean	Overall Std Dev
Prep for class	339	3.14	.575
Service to the college	339	2.87	.621
Administrator evaluation	339	2.50	.622
Self evaluation	339	2.48	.679
Service to students	339	2.40	.626
Student performance	339	2.33	.576
Alternative delivery formats	339	2.16	.700
Service to the community	339	2.01	.688
External evaluators	339	1.88	.591

^a172 CTE faculty + 167 transfer faculty responded to all items in this portion of the survey instrument.

Upon performing multivariate analyses, a few differences emerged between these two demographic groups of teaching faculty members. Using Wilks' Lambda it was determined that there were significant differences between CTE and transfer faculty members $F(1, 338) = 3.45, p < .001$, Wilks' $\lambda = .911$. Univariate analysis of variance (ANOVA) calculations were then performed as a follow-up to the significant findings from the MANOVA analysis on the nine scales for faculty performance to test for differences between the dependent variables (e.g. service to the college) with respect to the independent variables (primary teaching areas), and to determine which demographic group had the higher mean response for each of these nine scales. As shown in Table 27, the CTE and transfer faculty members differed in their ratings of the importance of four areas of performance.

Table 27

Primary Teaching Area: Between-Subjects Effects, ANOVA

Scale	Between-Subjects <i>F</i>	Significance**
Student performance	8.025*	.005
Service to the college	4.652*	.032
Alternative delivery formats	4.259*	.040
External evaluators	3.979*	.047
Service to the community	1.780	.183
Service to students	1.385	.240
Administrator evaluation	.818	.366
Self evaluation	.030	.863
Prep for class	.009	.924

* $p < .05$. **Bonferroni adjustment for multiple comparisons.

As shown in Table 28, for the items where there was a significant ANOVA finding, CTE faculty gave higher ratings to student performance and alternative delivery formats, while the transfer faculty rated external evaluators and service to the college as more important. The greatest differences between CTE and transfer faculty ratings (.16) were also for student performance and alternative delivery formats.

Table 28

Primary Teaching Area: Overall Means Comparisons

Item	CTE Faculty Mean	Transfer Faculty Mean	Mean Faculty Difference (Transfer – CTE)
Student performance	2.41	2.25	-0.16

Alternative delivery formats	2.24	2.08	-0.16
Service to the college	2.80	2.94	0.14
External evaluators	1.81	1.95	0.14

Years of full-time teaching. The length of full-time teaching experience was divided into three levels: (1) less than seven years, (2) seven to fifteen years, and (3) over fifteen years. The overall means and standard deviations are the same as in Table 26 since the two sample populations are identical. Consequently, preparation for class still received the highest overall rating ($M = 3.14$), followed by service to the college ($M = 2.87$), and administrator evaluation ($M = 2.50$). Service to the community ($M = 2.01$) and external evaluators ($M = 1.88$) were rated as the least important components of faculty evaluation.

Upon performing multivariate analyses, a few differences emerged between the three different years of experience demographic groups. Using Wilks' Lambda to investigate whether there were significant differences at the $p < .05$ level, it was determined that there were significant differences based on years of full-time teaching experience $F(2, 338) = 2.60, p < .001$, Wilks' $\lambda = .868$. Univariate analysis of variance (ANOVA) calculations were then performed as a follow-up to the significant findings from the MANOVA analysis on the nine scales for faculty performance to test for differences between the dependent variables (e.g. service to the college) with respect to the independent variables (years of full-time teaching experience), and to determine which demographic group had the higher mean response for each of these nine scales. As

shown in Table 29, significant differences emerged for three of the scales (administrator evaluation, preparation for class, and external evaluators).

Table 29

Years of Full-time Teaching Experience: Between-Subjects Effects, ANOVA

Scale	Between-Subjects <i>F</i>	Significance**
Administrator evaluation	8.688*	.000
Prep for class	3.549*	.030
External evaluators	3.454*	.033
Service to the community	1.680	.188
Alternative delivery formats	1.597	.204
Student performance	1.020	.362
Self evaluation	.850	.428
Service to the college	.774	.462
Service to students	.741	.477

* $p < .05$. **Bonferroni adjustment for multiple comparisons.

Table 30 shows the means for the three levels of teaching experience for the three significant ANOVA findings. Although the ANOVA calculations identified significant differences in the responses to the three faculty performance/service scales in Table 30 based on the faculty members' years of full-time teaching experience, further analysis was needed to pinpoint which particular demographic condition (<7, 7-15, or 15+ years of experience) differed significantly from the other conditions.

Table 30

Years of Full-Time Teaching Experience: Mean Comparisons

Item	N ^a	<7 Years Experience Mean	7 – 15 Years Experience Mean	15+ Years Experience Mean
Prep for class	339	3.05	3.10	3.25
Administrator evaluation	339	2.68	2.47	2.34
External evaluators	339	2.00	1.81	1.82

^a121 faculty with < 7 years experience + 93 faculty with 7-15 years experience + 125 faculty with 15+ years experience responded to all items in this portion of the survey instrument

As shown in Table 31, Scheffe post hoc analysis calculations determined that faculty members with more than 15 years of full-time teaching experience rated preparation for class higher than faculty members with less than 7 years of experience ($p = .011$). Faculty members with less than 7 years experience rated both (a) administrator evaluations higher than their colleagues with 7-15 years ($p = .019$) or more than 15 years ($p = .000$) of experience and (b) external evaluators higher than their colleagues with 7-15 years ($p = .030$) or more than 15 years ($p = .020$) of experience.

Table 31

Scheffe Post Hoc Analysis of Years of Full-Time Teaching Experience

Scale	Years of FT Experience (a)	Years of FT Experience (b)	(a – b) Mean Difference	Significance
Prep for class	<7	7 – 15	-.050	.546
	<7	15+	-.191	.011*
	7 – 15	15+	-.141	.079

Administrator evaluation	<7	7 – 15	.211	.019*
	<7	15+	.335	.000*
	7 – 15	15+	.125	.155
External evaluators	<7	7 – 15	.187	.030*
	<7	15+	.181	.020*
	7 – 15	15+	-.006	.946

* $p < .05$

Faculty gender differences. The final demographic comparison was for gender differences between full-time teaching faculty members in their responses to the nine faculty performance/service scales. Again, overall means were presented in Table 26 for all faculty respondents. Using Wilks' Lambda to investigate whether there were significant MANOVA differences at the $p < .05$ level between the mean responses to the faculty performance scales, it was determined that there were significant differences by gender $F(1, 338) = 2.10, p = .029$, Wilks' $\lambda = .944$. Univariate analysis of variance (ANOVA) calculations were then performed as a follow-up to the significant findings from the MANOVA analysis on the nine scales for faculty performance to test for differences between the dependent variables (e.g. service to the college) with respect to the independent variables (gender), and to determine which demographic group had the higher mean response for each of these nine scales. As shown in Table 32, the only area that showed a significant gender difference was with respect to the utilization of alternative delivery formats when teaching. Women ($M = 2.25$) placed a higher level of importance on this activity than did their male colleagues ($M = 2.05$).

Table 32

Gender: Between-Subjects Effects, ANOVA

Scale	<i>F</i>	Significance**
Alternative delivery formats	6.663*	.010
Service to the community	2.645	.105
Administrator evaluation	1.783	.183
Prep for class	1.768	.185
Service to the college	1.427	.233
Self evaluation	.651	.420
Student performance	.613	.434
External evaluators	.408	.524
Service to students	.046	.831

* $p < .05$. **Bonferroni adjustment for multiple comparisons.

No significant interactions. MANOVA calculations revealed no significant differences for any combinations of two or more of these three demographic variables (teaching area, years of experience, and gender). See Table 33 for a summary of the various MANOVA calculations related to demographic differences in the responses to the nine derived scales.

Table 33

Summary Wilks' λ Results for all Demographic MANOVA Calculations

Independent Variable(s)	λ	<i>F</i>	Significance
Primary role (faculty or administrator)	.853	8.78*	.000

Primary teaching area (CTE or general education)	.911	3.45*	.000
Years of full-time teaching (<7, 7-15, or 15+)	.868	2.60*	.000
Gender (female or male)	.944	2.10*	.029
Years of full-time teaching X gender	.919	1.54	.072
Primary teaching area X gender	.968	1.16	.324
Primary teaching area X years of full-time teaching	.955	.831	.664
Primary teaching area X years of full-time teaching X gender	.970	.540	.939

*p<.05.

Summary of Quantitative Results

Approximately 23% of the full-time teaching faculty members and 83% of the academic deans from 21 VCCS colleges responded to the on-line survey instrument. When quantitative responses were analyzed and demographic comparisons were made through the use of analysis of variance (MANOVA and ANOVA) calculations, several significant differences were found between the mean faculty and administrator responses. For the purposes and uses of faculty evaluation, administrators had higher means than the faculty members for (a) tying the results to professional development, (b) granting some discretion to the evaluators, (c) using the process for both formative and summative purposes, (d) requiring that faculty members receive two or more excellent evaluations for promotion eligibility, and (e) using the process primarily for formative purposes. In addition, whether using the faculty evaluation process for formative or summative purposes, administrators rated scholarly or creative activities and service significantly

higher than the faculty members. With respect to teaching performance, administrators placed greater importance on (a) operating a chat room or discussion board, (b) facilitating small group experiential learning events, (c) creating an on-line course, (d) grading examinations, and (e) developing course materials than the faculty members. Faculty members placed a greater importance on their teaching loads. Administrators placed greater importance on several contributing sources for the evaluation process: (a) dean evaluation, (b) systematic student evaluations, (c) direct supervisor's evaluation, (d) VP/Provost evaluation, (e) faculty compilation of a detailed portfolio, (f) committee evaluation, (g) student examination performance, (h) grade distributions, and (i) informal student opinions.

Factor analysis was performed on the survey items related to teaching performance, resulting in six factors: (a) student performance, (b) preparation for class, (c) self evaluation, (d) administrator evaluation, (e) alternative delivery methods of instruction, and (f) external evaluation. The survey items related to service performance were also evaluated using factor analysis and three factors emerged: (a) service to the college, (b) service to the community, and (c) service to students. These nine factors were then used as the dependent variables for examining differences between faculty and administrators as well as for evaluating differences between different faculty demographic groups.

Administrators responded with significantly higher mean values on six of the nine scales: (a) alternative delivery formats, (b) service to the college, (c) administrator evaluation, (d) service to students, (e) service to the community, and (f) student performance. Career and technical (CTE) faculty rated student performance and

alternative delivery formats higher than their transfer colleagues who, in turn, rated service to the college and external evaluators higher than the CTE faculty. Faculty members with more than 15 years of full-time teaching experience rated preparation for class higher than faculty members with less than seven years of experience and faculty members with less than seven years experience rated both administrator evaluations and external evaluators higher than their colleagues with seven or more years of experience. The only gender difference was for alternative delivery formats, where women had a significantly higher mean response than their male counterparts.

Qualitative Analysis

The survey instrument contained four open-ended questions to collect qualitative information on VCCS faculty perceptions of their evaluation processes. Two of the four questions addressed strengths and limitations, the third question asked about changes needed to local college plans, and the fourth question asked respondents to provide feedback on any areas of faculty evaluation they felt had not been adequately covered in the survey instrument. There were 302 responses to the question about strengths, 301 responses to the question pertaining to limitations, 290 responses to the question about what needs changing, and 95 responses to the request for additional comments.

Each set of responses was reviewed for the emergence of common themes through the process of content analysis – grouping responses into similar, non-overlapping categories of related themes, ideas, meanings or connotations (Stemler, 2001). To minimize bias in the analysis, emergent coding was employed to categorize each individual response into related groupings by utilizing the following procedure: (a) the author and a second researcher (a dean with 20+ years of experience with the VCCS

faculty evaluation process) independently reviewed the responses to each question to create a list of emergent themes, (b) the two investigators met to compare notes and to agree on a common list of emergent themes, (c) the common list was then used by the two investigators to independently review the responses in detail and to code each response into one of the identified themes, and (d) the coding of 20% the responses to each question was compared to check for the reliability of the coding. Cohen's Kappa was calculated for the coded responses. Stemler recommended agreement at the 95% level with a Cohen's Kappa of 0.8. Since the two investigators achieved at least 95% agreement on the 20% sample of responses, the principal investigator continued to code the remainder of the responses using the identified themes.

Research question 4. What do VCCS faculty members and administrators perceive to be the strengths and limitations of the current faculty evaluation process?

The investigators had a Cohen's Kappa value of 0.84 for the responses to the survey question related to the strengths of existing college faculty evaluation plans, and a Cohen's Kappa value of 0.86 for the responses to the survey question regarding limitations of their current faculty evaluation plans. These two areas will be discussed, below, as they relate to research question number four.

Strengths. The top emergent themes for current faculty evaluation plan strengths and their response frequency rates are presented in Table 34, which summarizes the content analysis findings based on the respondents' comments. Faculty and administrator responses were not analyzed separately due to the relatively low number of administrator responses. Table 34 shows the eight response themes as determined by the investigator

and co-evaluator. Only 4.9% of the responses did not fit into one of the eight identified themes.

Table 34

Response frequencies and percentages for emergent themes for the strengths of college faculty evaluation plans

Theme	# of Responses ^a	% Responses ^b
Use of multiple measures	57	18.7
Predictability or regularity of the process (including supportive administrators)	49	16.1
Formative feedback to faculty (including from students)	44	14.4
Faculty involvement/control of the process	33	10.8
Self-evaluation/reflection opportunity	30	9.8
Current plan has no strengths	30	9.8
Outcome-driven plan based on teaching performance - includes professional development/recognition	28	9.2
Interactions with both students and supervisor	19	6.2
Miscellaneous and non-responsive comments	15	4.9
TOTALS	305	99.9

^aTotal does not equal the number of individual respondents since several individuals gave multi-part responses. ^bTotal % responses may not equal 100 due to rounding of individual percentages

Three themes emerged as the greatest strengths cited from current faculty evaluation plans. The use of multiple measures, at 18.7%, was the most frequently mentioned plan strength. Some comments related to the strength of using multiple measures included, "The faculty are primarily evaluated based on self evaluations, peer

evaluations, student evaluations, and finally the Dean's evaluation. In an ideal setting, each of these are very good measures of performance. The strength is in the combination of all of them.” Another, response was, “They allow for diverse methods of demonstrating value to the college, not simply committee or community work, but personal/professional development. We don't have to pretend to be interested in something we're not just to please a committee/dean/president.” Another respondent noted the practical benefit of using multiple measures, “...so that different faculty with differing strengths can be fairly evaluated.”

The second most commonly identified strength was the predictability or regularity of the evaluation process, including supportive administrators (16.1%). As reported by one respondent, “My division has created a document that outlines exactly what will be part of the evaluation with points awarded for various activities. Faculty know what and where they have to contribute and have control of what goes into their evaluation. Minimal personal feelings/perceptions/ favoritisms involved.” Similarly, another faculty member wrote, “My dean allows me to provide input for my evaluation before she evaluates me, then discusses the results of her evaluation with me.” More straightforward comments cited the “The regularity of evaluations.” Perhaps the most telling comments related to this topic included responses similar to this statement, “Easy to get an excellent. This is good for promotion, but not good as a tool for faculty improvement.”

Formative feedback was the third most commonly reported strength (14.4%) and many faculty members valued the formative feedback they received from student evaluations. Comments related to the formative use of student evaluations of teaching included, “Student comments are always appreciated” and “Student evaluations are used

for self-improvement.” Similarly, another faculty member wrote, “Getting feedback from the students and figuring out what is constructive” was one strength of his/her faculty evaluation plan. A more reflective faculty member wrote, “I think it is always good to get feedback on your work performance...Another strength is the student evaluations, their opinion should be the one that matters the most.”

Faculty members also stated that the strengths of their evaluation plans included their own involvement in the process (10.8%) and the opportunity to include a self evaluation component in their evaluation documents (9.8%). Several faculty members wrote comments very much like this one respondent: “Faculty can manage the percentages of which category counts what percentage. Further, faculty control changes to the method of evaluation.” In other words, “Faculty have lots of input on the process.” Another faculty member went so far as to say, “The stipulation that faculty own the evaluation process; changes cannot be made without the vote of the Faculty Senate...The faculty member's leeway to write/format his/her own annual report. This MUST remain a faculty-centered/directed process.”

While some faculty members were happy about their “control” of the faculty evaluation process, others appreciated the opportunity to engage in self-reflection and for the use of this information in their evaluation. As one person wrote, “The narrative evaluations require faculty members to do an introspective look at his/her own teaching abilities and accomplishments as well as need for areas of improvement.” Or, as another faculty member wrote, it is “a time to reflect on our successes and failures.” Beyond simple self-reflection, the self-evaluation provided one respondent “a chance to think about, set and achieve goals and to talk with my supervisor.”

The final strength theme that elicited responses from almost 10% of the respondents was that their faculty evaluation plans were outcome-driven and based on their actual teaching performance (9.2%). Representative comments included, “The faculty evaluation is based on our jobs and not the extracurricular activities,” “It is heavily weighted in favor of our teaching ability,” and “The current evaluation process stresses the importance of teaching at the community college level.”

Limitations. The top emergent themes for current faculty evaluation plan limitations and their response frequency rates are presented in Table 35, which summarizes the content analysis of the respondents’ comments. As with the content analysis of plan strengths, faculty and administrator responses were not analyzed separately and Table 35 shows the six response themes as determined by the investigator and co-evaluator. While three themes emerged with high response rates, there was a drastic reduction in responses for any other themes, with 10.2% of the responses not fitting into any larger theme.

Table 35

Response frequencies and percentages for emergent themes for the limitations of college faculty evaluation plans

Theme	# of Responses ^a	% Responses ^b
Generally poor plan – not reflective of actual faculty duties, etc.	83	27.3
Heavy reliance on student input	79	26.0
Lack of objectivity - arbitrary or biased	71	23.4
Miscellaneous and non-responsive comments	31	10.2
Lack of multiple measures	27	8.9

No weaknesses reported	13	4.3
TOTALS	304	100.1

^aTotal does not equal the number of individual respondents since several individuals gave multi-part responses. ^bTotal % responses may not equal 100 due to rounding of individual percentages

Among the limitations cited in Table 35, the most common concern (27.3%) was that the faculty evaluation was simply a poor plan, with the most common complaints centered on the idea that the plan did not do a good job of measuring actual faculty responsibilities. As one faculty member summarized:

I think the whole process is a joke. It is an exercise in futility with many hours spent on the faculty's part and the administration's part writing a lot of "stuff" for what?? Faculty who are deserving of an excellent do not receive that rating, and faculty who do not deserve such a rating, get it. There is no requirement of "evidence" that the person has done what he/she says he/she has done. It does not take into account work load that some faculty have over others...The eval process needs to be meaningful.

Other comments on this topic included the statement, "The process does not foster open, honest communication. Supervisor and faculty approach it defensively, the former to make sure they are not sued or grieved, the latter to make sure they receive merit pay." There were several comments related to the idea that evaluation ratings were a political exercise with no connection to actual faculty performance. Statements included, "There is no meaningful evaluation process. It is a top-down administrative function" and "(t)oo check list oriented. Has been used as a tool of punishment by some deans."

Student evaluations were cited as a strong limitation (26.0%). Many individuals responded that student feedback carried too great a weight in their performance

evaluation, making this theme the second most commonly reported limitation. One consequence of a heavy reliance on any one measure, such as student evaluations, is that it does not provide a very well-rounded perspective of faculty performance. Out of the 24 items related to teaching performance that were presented in Tables 14 and 17, systematic student evaluations were rated as less important than direct supervisor's or dean's evaluation, but more important than self, peer, or external evaluation.

Some comments on the subject of limitations included, "Evaluations are based almost entirely on student evaluations of the instructors" and "The student evaluations are anonymous, have unaccountability, and are unreliable, and yet are used to evaluate faculty for the evaluation. On occasion, for example, one evaluation by students was reported in print for a class that I did not teach. I must add it was an excellent outcome score." Another faculty member referenced the "(e)xcessive numerical impact of student happiness ratings" at his/her institution. One faculty member was particularly opposed to heavy reliance on student evaluations, particularly when the student information came from the student response instrument in use at his/her college:

What instructor behaviors are we asking the student to evaluate? Determine that, then develop questions that test for those behaviors. This evaluation system does not do that. Do not ask any questions that ask "how much" a student agrees or disagrees with a statement; an instructor should not be rated by the level of "agreeness" a person has. I have no respect for our evaluation as written. It violates almost every protocol for developing legitimate surveys. I hate to think that such an evaluation is used to evaluate me as an instructor.

The third major limitation theme (23.4%) centered on the idea that the faculty evaluation process lacks objectivity, results in arbitrary ratings and/or is biased. As one faculty member wrote, “not everyone can be excellent but yet most are.” Another individual stated, “Sometimes subjective evaluation is not based on facts” as further explained in this response:

The limitations are in the fact that they are all subjective, and it is not usually an ideal setting. For example, the students may evaluate poorly based on personal dislike of the instructor; the instructor may inflate their self evaluation; the evaluators for the peer evaluations may not have much contact with the instructor or their students; and there is the subjectivity of the dean's evaluation.

To put it another way, “The limitations in faculty evaluation are: personality conflicts, pettiness, and the popularity contest elements” or “Your evaluation depends on the department you are in.” In particular, many faculty members expressed comments such as, “From my understanding, the evaluation can be very, very strongly influenced by your dean, so in my eyes, they are a little too subjective and reliant on one person. Fortunately I have a supportive and objective dean, but I have colleagues who do not.”

Individuals at colleges where there is a lack of multiple measures in the faculty evaluation process recognized this as a limitation of their plans (8.9%). Two, separate faculty members wrote, “There is not enough variety in the methods of evaluation used. It should include multiple measurements” and “(s)ources of input into the evaluation of individual faculty members is somewhat limited.” Another individual expanded on these ideas by stating, “Really MANY components go into our job and yet these are not

acknowledged or at least not measured...academic advising, supervisory functions, student recruitment, outreach, program promotion, etc, etc, etc.”

There were also many comments regarding faculty evaluation plan limitations that did not easily fit into any of the more common response categories (10.2%), yet some of the comments are worth mentioning due to repeated mention. Throughout the four qualitative response categories the topic of merit pay, or lack of merit pay, appeared periodically. Concerns were expressed that granting or withholding merit pay based on faculty evaluation ratings resulted in faulty processes. A typical comment on this topic was, “We are limited by too many ‘Excellents’ and not enough differentiation of results, due to merit pay attached to performance and lack of adequate pay raises for cost of living.” Some respondents felt that merit pay should be tied to the evaluation rating, “The faculty evaluation needs to be tied to something -- merit pay, continued employment, or something else...”

Another recurring limitation theme that appeared both in response to this question, and in the other open-ended survey questions, related to the amount of time devoted to the faculty evaluation process. At some colleges, the faculty evaluation process is so streamlined and simple, “You don't spend a lot of time you need for other tasks.” At other colleges, “evaluations take up entirely too much of the evaluator's time.” For a faculty member, not having to go through the process each year would be a relief, “Too time consuming to do a self-evaluation every year. Could be done every 3 years if satisfactory.” Administrators also would appreciate not having to evaluate every faculty member every year. “Insufficient time to prepare evaluations (two weeks for 30 faculty)” was one dean's comment on the subject.

Research question 5. What do VCCS faculty members and administrators suggest should be changed in the current faculty evaluation process?

The third open-ended question on the survey instrument was, “What needs to be changed with the faculty evaluation process at your college?” There were 290 responses to this question which were analyzed using the same content analysis method for determining emergent themes as for the preceding research question. The investigators had a Cohen’s kappa value of 0.82. The top emergent themes, and their response frequency rates, for this question are detailed in Table 36.

Table 36

Response frequencies for emergent themes for the suggested changes needed for college faculty evaluation plans

Theme	# of Responses ^a	% Responses ^b
Make the process more realistic/objective/meaningful with respect to actual faculty duties/activities	86	28.4
Need a more holistic approach with multiple measures	44	14.5
Need to simplify the process and make it useful	39	12.9
Reduce the emphasis on student evaluations	37	12.2
No changes needed	36	11.9
Miscellaneous and non-responsive comments	32	10.6
Make it more formative in nature	29	9.6
TOTALS	303	100.1

^aTotal does not equal the number of individual respondents since several individuals gave multi-part responses. ^bTotal % responses may not equal 100 due to rounding of individual percentages

By far, the most commonly suggested theme for changes (28.4%) included comments related to tying the evaluation process to actual faculty duties and for the process to have real consequences in terms of pay and/or retention of faculty. One respondent captured the essence of many other respondents when writing:

It's difficult for supervisors to make the call, but recognizing and rewarding top performers is a must. There is no incentive for continuous improvement except for professional pride. Essentially, you get what you pay for and you should pay for the best.

In addition to the concerns that the overall evaluation rating is not truly reflective of actual faculty performance, another respondent raised the issue of “merit pay” when commenting “The evals don't carry any weight. There is no merit money. Rewards are limited. Those who do the minimum to get by are rewarded exactly the same as those who do as much as they can.” There were additional comments centered on the idea that there should be real consequences tied to the evaluation rating, both positive and negative. As one person stated:

I think the main thing that needs to be changed at our institution is for the faculty evaluation process to actually mean something - to have some weight tied to it. As ours is currently written, it is nearly impossible to NOT score in the highest ranking making the whole process almost meaningless.

Other respondents added “Little is actually done to get ‘poor’ faculty to show improvement in teaching,” “Not closely enough tied to salary. Performers seem to get same ratings as non-performers,” and “It has never been taken seriously to root out incompetence.”

As discussed in the context of strengths, the use of multiple measures was the second-most mentioned area needing change (14.5%). As stated explicitly by one respondent “We need multiple forms of assessment including self evaluation. Our process includes student evaluation forms for fall semester only and dean evaluation. This is limited, not very comprehensive.” Others wrote “Use more of a portfolio approach which could include student evaluations as well as peer observations, course materials, and writings or art work,” and “I think that the evaluation process could be more holistic.” One self-identified dean provided this comment within the body of a much longer response “I agree with the faculty that a multi-method, multi-measures, approach is the best.”

The third-most common theme related to a desire to see the faculty evaluation process simplified and made more useful (12.9%). Several comments stated that the current faculty evaluation processes in place at their institutions was too time-consuming “The process is very lengthy, time consuming and cumbersome. The evaluation should be streamlined where possible” and “Requires TOO much time that could be spent better working with students or curriculum.” Other comments included the desire to streamline evaluations “to meet the job description” and “It must have some meaning or it is just busy work.” Suggestions for streamlining the process included “I think a checklist would suffice. We could submit additional materials as support” and “If faculty has a five year appointment, then only do a major evaluation once every five years.”

While the extensive use/reliance on student evaluations was the second-most cited limitation of current faculty evaluation plans, it was only the fourth-most common theme for change (12.2%). Comments ranged from those concerned with the validity of the

student evaluation instruments “I think the student evaluations are poorly constructed and the information collected is not as helpful as it could be” to those who feel “Student reaction to instruction is used as a weapon in my division. Deans may cherry pick negative comments to include in the teaching evaluation.” One of the most constructive comments related to student evaluations of teaching faculty suggested the “creation of a much better student evaluation instrument, ideally one created outside the college and norm referenced to similar institutions.”

The final theme cited as needing change was that the process should be more formative in nature (9.6%) with an emphasis on professional development for continuous improvement. As summarized by one respondent “It needs to be more formative and less summative. It needs to have consequences and provide a path to improvement.” Others had stronger opinions on this topic that illustrate the contentious nature of the faculty evaluation process, while essentially agreeing that the process should be used for faculty improvement. One respondent stated the need to “Provide proper training of supervisors & deans to conduct fair evaluations” while another individual stated that “Faculty need to get off their dead butts and get serious about what is important rather than what is easy to count. It would also be handy if they trusted deans more -- it's pretty dumb not to trust the folks we have and are likely to acquire.” Another individual took the larger view when responding “this is bigger than the individual colleges, system wide focus needs to be on assessing teaching & learning and continuous improvement, consider SACS requirements.”

Other Feedback

Survey participants were also given an opportunity to address any remaining issues they had with faculty evaluations. The fourth and final open-ended opportunity to provide feedback on the survey instrument was: “Please use the following space to address any issues related to faculty evaluation that you feel were not adequately covered in this survey instrument.” There were 95 responses to this question which were analyzed using the same content analysis method for determining emergent themes as described for the preceding research questions. The investigators had a Cohen’s kappa value of 0.84 and the top emergent themes, with their response frequency rates, as detailed in Table 37.

Table 37

Response frequencies for emergent themes for other areas of concern related to the college faculty evaluation plans

Theme	# of Responses ^a	% Responses ^b
Unbiased evaluation with standardized criteria and true consequences	27	26.7
Miscellaneous and non-responsive comments	20	19.8
No other issues	13	12.9
Multiple measures to evaluate the “whole picture”	13	12.9
Less emphasis on student evaluations/outcomes	11	10.9
Training for evaluators	7	6.9
More frequent faculty/administrator interactions, formative approach with professional development	6	5.9
More emphasis on student evaluations/outcomes	4	3.9
TOTALS	101	99.9

^aTotal does not equal the number of individual respondents since several individuals gave multi-part responses. ^bTotal % responses may not equal 100 due to rounding of individual percentages

The most common theme (26.7%) related to the desire to have an unbiased evaluation process that utilizes standardized criteria and has true consequences related to faculty pay and retention. As one respondent wrote “This survey instrument only assesses the current form of performance evaluation. If all faculty (members were) held to the same standards and work load then we could begin to rate faculty by the same evaluation system.” Many others wrote that their current process does not provide unbiased ratings, as explained by this comment:

The survey does not address the question of who determines the definition of a “good” faculty member. The evaluation process at (my college) is hierarchical and heavily politicized. It is beset with egos, quirks, competitiveness, and convoluted power struggles. It has no relationship to quality.

In addition, another respondent addressed the issue of trust between the faculty member and their evaluator:

The survey did not deal with the issues of training or trust. No evaluation process will be effective if the evaluators are not trained to evaluate objective performance criteria and if the faculty do not trust administrators to evaluate fairly, consistently, and objectively.

Others added, at their colleges, there is no true comparison of relative faculty strengths and “everyone receives an excellent, which makes it hard to recognize those who are truly excellent, and the fact that really bad faculty, who need to go, rarely receive worse than a good.” One other respondent expressed another viewpoint: “Again, we are not actually measuring what is important. We measure mindless crap like committee service

and community involvement and don't look at our actual job which is to get the students to meet their goals!"

While there were a large number of miscellaneous or non-responsive comments (19.8% of the responses failed to "fit" into any larger theme), and many respondents wrote to say that they did not have any other issues (12.9%), the next largest theme of responses was for the use of multiple measures (12.9%), which has already been adequately explored in other areas of the data analysis. However, one interesting comment made in response to this prompt was:

I believe that for the summative teaching evaluations scholarly activities & professional recognition should be considered as "extra credit", i.e., if present, used to enhance one's score but if absent not lowering one's score. I think that community service should be considered in the same way. Many of us, myself included, provide "community service" by going way above and beyond the call of duty or remuneration in service to the students in our classes. Should the evaluation process encourage cutting back on this "service" in favor of more traditional community service?

Student evaluations also re-emerged in this section but the responses were divided between those who oppose and those who support use of student evaluations in the faculty evaluation process. Those who advocate a reduction in the use of, or reliance on, student evaluations and outcomes outnumbered, by a ratio of nearly three to one (10.9% to 3.9%), those who favor greater use of student evaluations. One comment in favor of reducing reliance on student evaluations was:

Student opinion has taken on much too much weight nationally in evaluation of professors--the pendulum NEEDS to start swinging the other way. I do consistently get positive comments from students, so this is not sour grapes. My concern is that students are no more qualified to evaluate professors' job performance than we are to evaluate theirs in their chosen fields, and when student opinion is in any way tied to faculty evaluation, compensation and in some cases even retention decisions . . . this is a national problem and will, I hope, be a helpful comment for your research.

In counterpoint to the above view, another respondent wrote:

Personally, I do not like the idea of a governing body of campus administrators or committee evaluating a faculty member's teaching performance. In my opinion, the students are the ones who should evaluate their teacher's performance in the classroom, as they are the ones who are directly affected by the teacher's performance. Though I understand the need to maintain quality of instruction, I have a problem with a campus official(s) (who likely have not set foot in a classroom as a teacher for many years) offering suggestions of improvement in teaching technique. I also equally dislike the idea of peer evaluations (by other faculty), as there are many individual differences in teaching technique that can bias these evaluations of what constitutes "better teaching". Again, as a teacher, my foremost commitment is to my students. Therefore, I believe that THEY are the ones in the better position to evaluate the quality of instruction and teaching performance. I further believe that it is THE STUDENTS whose opinions really matter with regard to evaluation of teaching performance.

Training for evaluators emerged as one of the lesser themes (6.9%) for this part of the survey, yet it is in keeping with comments expressed in response to earlier questions on the survey instrument. Some comments that did not fit into any of the larger themes that emerged for either strengths or limitations, but appeared as responses to both areas included comments on the large role that deans play in the faculty evaluation process. Drawing from comments made in response to the first three open-ended survey questions, one individual stated that a strength of his/her plan was “These depend almost entirely on the administrator who is conducting the evaluation” while a response to the limitations question was “Deans are not consistent. If faculty had a choice of which division dean they'd prefer, we'd have some empty divisions. There is too much favoritism in play with regard to faculty evals.” In other words “Each dean may implement the evaluation process in a different manner.” Considering these views, it was not surprising that several respondents to the “other areas of concern” question suggested that evaluators should receive training. One comment addressed the issue directly:

The survey does not really address the potentially subjective nature of the evaluator. Evaluations should not be strictly number crunching, but there is too much room for the evaluator's subjective measuring of the components in the evaluation other than student evaluations. It is not clear how those doing the evaluations have been trained to serve that role; I don't see anything in your survey about how one determines if the evaluator is in a position to perform a viable, meaningful evaluation.

Returning to the “theme” of miscellaneous comments, there were a number of interesting statements such as “We desperately need VCCS-wide job descriptions for

everyone, esp. faculty and much better guidelines on the evaluation process.” Other respondents asked their own questions in response, such as “Issues to consider: How political do you feel faculty evaluations are in your college? Are annual evaluations true reflections of faculty's performance?” and “How will the VCCS use your results--or will they ignore them because they don't want to "fiddle" with formative aspects?”

Summary of Qualitative Results

From approximately 1,000 responses to the four open-ended questions on the survey instrument, several common response themes emerged through the process of content analysis. Comparing the faculty and administrator responses to questions regarding current faculty evaluation plan strengths, limitations, and suggested changes, the qualitative responses indicated some qualities that good faculty evaluation plans possess, and poor plans do not possess: (a) the use of multiple measures/sources of input, (b) based on actual faculty duties, (c) student evaluations are utilized for formative feedback, but not as the sole source of information, (d) the plan should be objective and based on valid, measureable criteria, and (e) connected to professional development for both faculty and administrative evaluators.

CHAPTER 5

DISCUSSION

The purpose of this study was to address an area of critical importance to both Virginia Community College System (VCCS) faculty members and their supervisors – faculty evaluation. While the topic of faculty evaluation is certainly not new (Remmers, 1930), it is still a contentious topic. Searching the Education Resources Information Center (ERIC) database using the terms “evaluation” and “college faculty” produces a list of over 5,800 articles published since 1966. This chapter begins with a discussion of quantitative results that pertain to the (a) purpose of faculty evaluation, (b) composition of typical faculty evaluation plans, (c) data sources used in determining the faculty rating, (d) areas of faculty responsibility that may be factored into the faculty rating, and (e) demographic differences regarding the components used in a faculty evaluation plan. Qualitative results regarding the strengths, limitations, and suggested changes to VCCS faculty evaluation are also discussed. Significant findings are discussed as they relate to the literature and to practical applications for the VCCS faculty evaluation process and/or individual college plans. Limitations of the study are presented, directions for further study on this topic are suggested, and implications for VCCS policy and practice are addressed.

Quantitative Results

The Purpose of Faculty Evaluation

Evaluation of teaching faculty members is often a controversial process, primarily because a single evaluation process serves two often conflicting purposes (a) to provide *formative* feedback to faculty members so that they can improve their performance and

(b) to provide *summative* feedback to assist administrators in making personnel decisions related to retention, promotion, or dismissal of teaching faculty members. Participants in the present study responded to the prompt “both formative and summative processes” with the highest mean out of seven proposed uses for faculty evaluation. When considered separately, using faculty evaluation for summative purposes received the lowest mean response by survey respondents. Morris (1997) suggested faculty improvement (formative evaluation) and institutional accountability (summative evaluation) were two goals met by a single faculty evaluation process. To conform to accreditation criteria, Texas community colleges, by law, have faculty evaluation plans that are both formative and summative (Campion, Mason, & Erdman, 2000). The present survey respondents endorsed both formative and summative uses throughout their quantitative and qualitative responses. A properly constructed and conducted faculty evaluation process can meet these two goals, but this is not a simple process: it requires a combined effort on the parts of both faculty members and their administrative supervisors (Arreola, 2007; Lee, 2006).

The underlying unspoken goal of the faculty evaluation process is improvement in the quality of student education (McGee, 1996; South Texas College, 2004). If indeed the underlying reasons for conducting the faculty evaluation process are improved teaching effectiveness and quality of student education, then using the results of the faculty evaluation process for both formative and summative purposes makes sense. A formative process provides faculty members with specific information from students, who are the most frequent observers of teaching (Scriven, 2005) and faculty members can use this information to improve their teaching performance. Qualitative responses to the present

study supported formative feedback from students as one of the major strengths of their current evaluation plans. In turn, a summative process can provide administrators with more global information from students (Cashin & Downey, 1992), and other contributors, to assist in the decision processes of retention and promotion of “good” faculty or the dismissal of “not good” faculty members.

The literature suggests that faculty members may reasonably expect the evaluation process to be formative in nature for their personal use, while administrators are more likely to use the results of the faculty evaluation process in a summative fashion to inform decisions related to personnel matters such as raises and renewal of contracts, or for promotion and tenure decisions (Campion, Mason & Erdman, 2000; Morris, 1997; Worcester, 1993). Contrary to this expectation, the current results suggest that VCCS administrators had significantly higher mean responses for using faculty evaluation for formative and professional development purposes than did the faculty themselves. Additionally, administrators had a higher mean response with respect to having some discretion in assigning the final evaluation rating. Perhaps, as Seldin reported (1999), this is because administrators have access to all of the other sources of data that go into the faculty evaluation process and have a better over-all view of faculty performance. Qualitative faculty responses, discussed later in this chapter, also addressed the topic of administrator bias in the evaluation process.

With respect to formative and summative evaluation purposes, teaching rose to the top of the list as clearly the most important element for inclusion in the faculty evaluation plan, as expected. All other faculty activities had much lower means than teaching. When asked additional questions, items related to teaching in this study

received the highest mean values from both faculty and administrators. McGee (1996) reported in his study of 247 community colleges, the top factor contributing to faculty evaluation was classroom teaching. Professional recognition and scholarly or creative activities did not make McGee's list of the top five contributing factors to faculty evaluation and they were also the bottom two responses to this study; well below teaching, personal attributes, and service. These results support the contention that community college teaching faculty are very focused on teaching, not on professional recognition and scholarly or creative activities, unlike the faculty at senior institutions. This finding points out the contradiction between community college faculty and the findings of authors who restricted their studies to faculty at four year institutions.

Composition of the Faculty Evaluation Plan

Since teaching is the primary activity undertaken by community college teaching faculty, on what other measures should their performance be rated? Indeed, a major stumbling block that impedes the development of a useful faculty evaluation plan according to Centra (1993) and Seldin (1999) is the lack of agreement on the duties or activities in which a faculty member engages that should be included in the evaluation process. As might be expected, for six of the 11 teaching activities included in the survey instrument, faculty and administrators differed significantly in their responses regarding inclusion of these activities in the faculty evaluation plan. While the overall mean responses were highest for "traditional" teaching activities such as developing course materials, exams, and developing lecture materials and lowest for "non-traditional" activities such as operating a chat room or discussion board, facilitating small group experiential learning activities or creating an on-line course, these non-traditional

activities received significantly stronger administrator response means compared with the faculty response means. In other words, administrators favored innovation and alternative teaching methods while the faculty respondents preferred traditional lecture related teaching activities.

These results support Arreola's (2007) contention that the actual composition of faculty evaluation plans should be agreed upon by both faculty and administrators. Faculty duties are so varied that not all activities should be equally expected of all faculty regardless of their academic discipline and work assignments. Otherwise, the process will continue to be contentious and full of the distrust mentioned in some of the qualitative responses.

Data Sources Used in Determining the Faculty Rating

A faculty evaluation plan that is meaningful to faculty members and their supervisors requires strong agreement on the data sources to include in the plan. McGee (1996) found that the three most common sources of information utilized in faculty evaluation plans were student (92%), supervisor (84%) and self (72%) evaluations. In 2002, Paulsen suggested the three most common sources of data were student ratings, peer evaluations and self-evaluations/portfolios. South Texas College (2004) revised their faculty evaluation plan process and determined that their faculty evaluations would be composed of instructor self-evaluations, classroom observation by supervisor or lead instructor, student evaluations, and a summary administrator review. Arreola (2007) noted that it is important to specify the proportional weight each source of information has on the total evaluation.

While there is not universal agreement, the following major data sources appear in most faculty evaluation plans (a) student ratings, (b) peer evaluations, (c) self evaluations, and (d) supervisor evaluation. As several authors made clear (Arreola, 2007; Centra, 1993; Seldin, 1999; Theall, 2005), utilizing multiple sources of information to provide a balanced faculty evaluation plan is very important. The literature is highly variable with respect to the value of each of these components of faculty evaluation, so each must be considered with the proviso that there is no universal agreement as to the validity or reliability of most evaluative instruments. The literature on faculty evaluation is so rich that one may find several articles to support nearly any position (pro or con) one wants to take on any aspect of the process, particularly with respect to the use of student evaluations. In alignment with the literature, faculty and administrators disagreed significantly with respect to the utilization of nine of the 13 potential data sources contained in the survey instrument.

Student evaluations. Not surprisingly, of the nine data sources for faculty evaluation where faculty and administrators had significantly different responses, student evaluations tied (with dean evaluation) for the greatest difference. This is in strong support of the literature where some of the faculty concerns relate to the inclusion of student evaluation in the faculty evaluation process (Wolfer & Johnson, 2003), the validity of student evaluations (Langbein, 1994; Yunker & Yunker, 2003), and the reliability of student evaluations (Wright, 2006). Other studies questioned whether student evaluations were truly reflective of the quality of instruction received by students or, rather, more a reflection of expected course grades (Millea & Grimes, 2002), personality of the instructor (Russell & Gadberry, 2000), gender of the faculty member

(Basow & Silberg, 1987) or students (Germain & Scandura, 2005), level of the course (Ory, 2001), degree of course difficulty (Addison, Best & Warrington, 2006), class size (McPherson, 2006), or student preparation/motivation levels (Davidovich & Soen, 2006). Nuhfer (2005) cautioned that no single measure, such as student satisfaction ratings, could adequately capture or describe a complex activity such as teaching. Nuhfer went on to state that “student ratings alone cannot capture ‘good teaching,’ prove that learning occurred, or serve to show outcomes were met” (p. 14). Therefore, a comprehensive faculty evaluation system should not rely on student evaluations alone. Several of these concerns were also expressed by VCCS personnel in their qualitative responses to the survey instrument.

Peer evaluations. Peer evaluations are less well researched, but there is evidence that this form of evaluation is growing (Osborne, 1998). Yon, Burnap, and Kohut (2002) concluded that peer observation reports could play an important part in the evaluation of teaching. Regarding the use of peer evaluations in the VCCS faculty evaluation process, faculty and administrators agreed that this data source was less important than supervisor, student, and self evaluations.

Self-evaluations. Self-evaluation may take the form of a simple checklist, a narrative summary, or an extensive portfolio. While a checklist has the advantage of simplicity – both for the faculty member to complete and for the evaluator to assign a rating - there is no evidence to support the validity of a faculty member’s self-ratings. The most comprehensive form of self-evaluation is when the faculty member creates a portfolio; and portfolios have been used for conducting both formative and summative evaluation of faculty. Melland and Volden defined a portfolio as a “compilation of

carefully selected materials reflective of one's teaching activities, presented in an organized manner by an individual faculty member" (1996, ¶2). While portfolios do offer teaching faculty members the opportunity to provide evidence of their professional activities that are not captured by student evaluation questionnaires, etc., they also may wind up being so complex and comprehensive that they become impossible to evaluate.

VCCS faculty and administrators disagreed on the inclusion of detailed faculty portfolios, perhaps due to the complexity of the portfolio evaluation process as stated by Theall (2005). This result corroborated Sain's (2008) study on the use of portfolios in faculty evaluation in the North Carolina Community College System, which found that administrators and faculty members differed in their view of the value of portfolios. Both faculty and administrators expressed concerns over the amount of time involved in utilizing the portfolio process (Sain). VCCS survey participants had greater agreement on the use of narrative reports and/or simple checklists for faculty self-evaluation.

Administrative evaluations. Very little information was readily available in the literature concerning administrative evaluation as a separate component of the faculty evaluation process. Perhaps administrative reliance on second-hand data fosters a sense of distrust between faculty and administrators. Redmon (1999) stated that community college faculty members, "generally share a belief that administrators should be more willing to share resources and power, allow for creative growth and development in teaching, and allow for greater adaptability in showcasing their professional growth" (p. 57). Also, Paulsen (2002) stated that "clarifying the expectations that institutions and departments have for their faculty and that faculty have for their own performance are central to a successful faculty evaluation system" (p.5). At many institutions, this sharing

of information between administrators and faculty members is woefully lacking (Seldin, 1999; Worcester, 1993).

With regard to administrative sources of data, as noted in the literature, it was no surprise that VCCS faculty and administrators disagreed on the importance of input from deans and direct supervisors. However, the two groups agreed that these data sources were the most important, closely followed by student evaluations. Although they had significantly different means, both faculty and administrators were less inclined to include evaluation by the VP/Provost or an external committee than they were to include student input. In summary, it is important that a comprehensive faculty evaluation process should allow for multiple data sources and the collection of relevant data throughout the evaluation cycle (Arreola, 2007; Lee, 2006).

Areas of Faculty Responsibility Factored into the Faculty Rating

Another reason that many faculty members view any faculty evaluation process as flawed is they have received little or no training for many of the tasks upon which they are evaluated (Arreola, 2007). Indeed, Milton and Shoben (1968) noted that college teaching is one of the only professions in the world where people with no specific training are hired to perform a complex task. Study faculty participants rated “non-traditional” teaching activities such as creating an on-line course, operating a chat room or discussion board, or facilitating small group experiential learning events at the bottom of the list of important teaching performance activities to include in their evaluation plans.

Faculty members are typically hired based on their subject matter expertise and/or research skills with little regard for whether or not they are trained on how to teach,

develop academic programs, construct and evaluate exams for students, etc. Regardless of whether or not there were significant differences between the faculty and administrator mean responses, VCCS respondents collectively rated preparation for class, service to the college, and service to students higher than student performance. Many of these responsibilities are ones for which the faculty members have very little preparation. Therefore, when a faculty member is rated low in any of these areas, she should receive professional development designed to provide the missing training. This leads us to Scriven's comment, the "implication is that the evaluation of faculty performance must be linked with institutional programs that support professional development as a necessary element in improving overall institutional performance" (2005, p. 9).

Demographic Differences in Responses

In addition to comparing the results of this study to the existing literature, this study was designed to contribute new information to the literature with respect to demographic differences in faculty and administrator views on the faculty evaluation process and on the activities and data sources that should be included in the process. It was logical to assume, based on Arreola's (2007) assertions that faculty evaluation plans should be individualized to suit differing individuals, departments, etc., that different groups of faculty members would place differing values on the use of certain components of the faculty evaluation plan, or might differ on which sources of data they valued. The scales developed by factor analysis of the teaching and service performance responses were used as the basis for conducting comparisons between various demographic subgroups of faculty respondents.

Primary teaching area. Analysis of variance calculations revealed that transfer faculty placed significantly greater importance on external evaluation, and service to the college than did their career and technical (CTE) colleagues. On the other hand, CTE faculty rated student performance and alternative delivery formats higher than their transfer colleagues. Transfer faculty understand that high student grade point averages are critical for successful transfer, so perhaps they feel that students should bear the burden of earning good grades. CTE faculty may be concerned that their students are able to apply their knowledge in the field, so they place a greater emphasis on student performance. The one area of difference that was contrary to expectations related to the higher rating of external evaluation by transfer faculty. In the VCCS, CTE faculty members meet with their advisory board members on a regular basis to discuss how well their programs prepare students for the workplace, and to make necessary adjustments. External evaluation is a standard practice for CTE faculty, so the higher rating by transfer faculty was unexpected. The literature is silent with respect to differences between transfer and CTE faculty views on faculty evaluation.

Years of full-time teaching experience. Faculty members with varying years of full-time teaching experience differed in their responses to administrator evaluation, external evaluation, and preparation for class. Faculty with less than seven years experience placed greater importance on these activities than their more experienced colleagues. These results compare favorably with Yao and Grady's similar findings that faculty members with more experience paid less attention to student evaluations (2006). Junior faculty members are eager to receive formative feedback in order to improve their performance while more senior faculty members are more set in their ways and less

interested in receiving formative feedback (Yao & Grady, 2006). It is, therefore, reasonable to expect that if more experienced faculty members place less value on student feedback than their less experienced colleagues they would also place less value on feedback from any other data sources internal or external to the college.

Gender. The only significant difference found between genders was for the use of alternative delivery formats of instruction, with females placing greater importance on this activity than their male counterparts. Literature on the effects of gender on student evaluations has been inconclusive (Ory, 2001) with respect to the ratings based on either the faculty members' or the students' genders, but nothing in the literature indicated any differences in instructional delivery preference by gender, and this result was completely unexpected. As noted earlier, alternative or "non-traditional" teaching activities received the lowest ratings among the teaching activities covered in this study and, at senior institutions, female faculty members tend to be assigned lower status assignments (Myers, 2008). Perhaps this gender difference is simply a reflection of the types of teaching assignments female faculty members are given, so women are more receptive to the inclusion of alternative forms of instruction in their evaluations. If this is the case, this gender difference result supports Stryker and Serpe's (1982) "identity-behavior" link. There were no significant findings for any of the interactions of these three demographic factors (primary teaching area, years of full-time teaching experience, and gender).

Qualitative Results

Strengths of the VCCS Faculty Evaluation Process

In agreement with the recommendations of Arreola (2007) and Lee (2006), the use of multiple measures was the most frequently mentioned plan strength. The

predictability of the process, including the awareness that the respondents had supportive administrators was the second-most cited strength. As noted by Arreola (2007), a faculty evaluation plan that is meaningful for both faculty members and their supervisors requires strong agreement on the data sources, component areas, and specific elements to include in the plan. The use of formative feedback, including the use of student evaluations of teaching, was the third-most cited strength. A formative process provides faculty members with specific information from students, who are the most frequent observers of teaching (Scriven, 2005) and faculty members can use this information to improve their teaching performance.

Limitations of the VCCS Faculty Evaluation Process

The most commonly cited limitation of current VCCS faculty evaluation plans was that the plans were generally poor with little or no relation to actual faculty duties. This finding strongly supports Arreola's statements that "the validity of any form is a function of the degree to which the form measures those aspects of faculty performance that faculty believe to be important to measure in the first place" (2007, p. 1), and "beginning with actual faculty performances provides us with a more accurate and complete definition of the roles faculty play as they pursue their professional responsibilities within the institution" (p. 3). As Centra (1993) wrote "Not everything that counts can be counted and not everything that can be counted counts" (p. 176).

As discussed above, it was anticipated from the voluminous literature on the use of student evaluations that this would be the most commonly reported limitation of existing VCCS faculty evaluation plans. However, while many faculty members responded that student feedback carried too great a weight in their performance

evaluation, making this theme the second most commonly reported limitation, others valued the formative feedback they received from student evaluations. Hobson and Talbot stated “Despite discrepancies in opinions and research findings on the validity of student rating, it is essential for faculty to understand that student evaluations are – and probably will continue to be – the primary institutional measure of their teaching effectiveness” (2001, ¶7). Many VCCS faculty and administrators apparently realize that the formative information available through student evaluations is meaningful, but student evaluations should not be the only data source utilized in the faculty evaluation process.

Another limitation that emerged from this study referred to plans that lacked objectivity and were administered in an arbitrary or biased fashion. In other words, these plans lacked objective, standardized criteria for measuring faculty performance. With no agreed-upon standards, it is impossible to arrive at objective faculty ratings (Arreola, 2007). This also speaks to the lack of trust between faculty and administrators mentioned by Redmon (1999) and several survey respondents.

Suggested Changes to the VCCS Faculty Evaluation Process

By far, the most commonly suggested themes for change related to the desire to have an unbiased evaluation process that utilizes standardized criteria tied to actual faculty duties as recommended by Arreola (2007) and for the process to have real consequences in terms of pay and/or retention of faculty. As stated by Arreola, “a true merit pay program is to recognize and reward past meritorious performance and to encourage future meritorious performance” (2007, p. 84). Currently in the VCCS, merit

pay is not objectively based on merit at most colleges, but serves as an across the board cost of living increase for faculty.

Limitations

While the survey population consisted of all academic deans and full-time teaching faculty members in the VCCS, individuals who had an interest in the topic were most likely to complete the survey. Therefore, the study group population consisted entirely of self-selected individuals. Due to the time required for the internal approval processes at some of the larger colleges in the VCCS, the survey was not distributed to faculty and deans at those colleges until after the end of the spring semester, a time when many full-time teaching faculty members were on their summer vacations and unavailable to participate in the study. The response rates at these colleges may have been negatively affected. It was also disappointing that two of the college presidents declined the opportunity to participate in the study, thereby preventing this study from truly being a system-wide study.

However, due to the wide-spread interest in the topic of faculty evaluation in the VCCS, it is probable that responses received were representative of the entire spectrum of administrative and teaching faculty opinions. Based on the variation in responses received, it was apparent that the survey captured information ranging from those opposed to faculty evaluations in general, to those who hoped to see a fair, objective process developed and adopted. Although the response rate for academic deans was very high (82.7%), the relatively small sample size (67 individuals) may have contributed to the lack of significant differences between administrator demographic groups. Also, the large ratio of faculty to administrator respondents (6:1) may have lead to statistically

significant differences appearing in the results that were primarily a difference due to the discrepancy in sample size. As always with a self-selecting sample population, extrapolation of results across the entire VCCS should be done with care.

This study involved only full-time Virginia community college teaching faculty members and the administrators who evaluate them. Therefore, the results may not confidently be generalized to be representative of the opinions/preferences of professional faculty members (counselors, coordinators and librarians), part-time (adjunct) teaching faculty members, or four-year teaching faculty members inside of Virginia, nor any teaching faculty members in other states.

A separate type of limitation associated with survey methodology is self-report; or social desirability. In addition, since participation in this survey was voluntary with no tangible reward for participation other than a chance to win a gift certificate, one must be aware that individuals who chose to respond tend to have strong reasons for responding – either in favor of or in opposition to the topic of the survey. As was established at the outset of this dissertation, faculty evaluation is a controversial topic, so strong opinions were expected to be expressed. However, assurances of maintaining respondent anonymity encouraged response candor and minimized social desirability for all participants.

Although some of the results from this study touched on student performance, the survey instrument did not directly ask VCCS faculty and administrators about tying faculty evaluation rating to student performance. With increased national attention on performance-based funding (Jaquette, 2006), perhaps some explicit questions should have been included, but the VCCS does not have a history of basing faculty evaluation

ratings on student performance. Jenkins (2006) discovered that it is difficult to document that student learning has occurred and stated: “One major obstacle to formulating a policy was describing the outcomes as well as developing activities to assist students in achieving those outcomes using appropriate assessment methods” (p. 4). Such considerations were outside of the scope of this study.

Directions for Further Study

Students and faculty members at community colleges are tremendously understudied populations in general. Considering that roughly 43% of all US (AACC, 2009), and 52% of Virginia’s (State Council, 2010), undergraduate students are enrolled at community colleges and receiving most of their instruction from full-time teaching faculty members, community college faculty are poorly represented in the literature. This study should be replicated throughout the country at representative institutions of the 1,173 US community colleges (AACC, 2009). Virginia’s 23 community colleges are part of a public system. It might be illuminating to replicate this study at some of the 155 private and 31 tribal community colleges to compare the findings for common and divergent results.

In addition to contributing to the body of literature on faculty evaluation in public community colleges, this study also examined differing views based on respondent demographic profiles. However, the study only scratched the surface of demographic differences with respect to faculty evaluations – an area that is sadly lacking in the literature. Additional studies related to faculty and administrator differences with respect to faculty evaluation should be conducted. Similarly, there should be more large-scale studies to compare differing demographic group responses to faculty evaluation.

Demographic variables could include gender, subjects taught, lecture versus laboratory instructors, years of experience, level of courses taught (freshman, sophomore, etc), type of institution (two-year, residential four-year, commuter four-year), average class size taught, or whether the faculty are teaching at liberal arts or research institutions. One size does not fit all with respect to faculty evaluations (Arreola, 2007), and the more that is known about the particular faculty group being evaluated, the more useful the process.

Survey Instrument as a Tool for Future Research

Since the survey instrument developed for this study has been validated through the processes of expert review, revision, and factor analysis, and reliability has been supported through moderately strong Cronbach's alpha coefficients, it would be instructive for others to utilize this survey instrument for assessing faculty and administrator responses at other community colleges for comparison with these VCCS results. Due to the small number of survey participants from several of the colleges participating in this study, college-specific differences were not calculated. Individual colleges would be well advised to conduct their own internal studies to assess local views on faculty evaluation.

Implications for Practice

There are three separate ways in which the results of this study could be utilized by the VCCS to improve the faculty evaluation process: (1) revision or creation of some VCCS policies at the system level, (2) college review and adjustment of current faculty evaluation plans, and (3) implementation of professional development opportunities for both faculty and administrators to improve the faculty evaluation process. These three potential uses, and the implications for using the results of this study are presented below.

VCCS Policy Revisions

The *VCCS Policy Manual* (VCCS, 2010a) should be examined for revision and/or the development of new policies. College plans should be carefully reviewed for compliance with Sections 3.5 (Faculty Responsibilities) and 3.6 (Faculty Evaluations) of the *VCCS Policy Manual*, see Appendix E for current policy language. There is currently no connection in the *VCCS Policy Manual* between suggested faculty responsibilities (Section 3.5) and expected faculty performance (Section 3.6). Specific job expectations differ by position and, therefore, different activities should receive different relative weightings in individual faculty members' evaluation plans. It has been suggested that detailed faculty position descriptions should be developed that contain job expectations, and these faculty position descriptions should be tied to the faculty evaluation process as the baseline expectation for earning a "good" evaluation rating. Participants in this study expressed the strong desire to have a meaningful faculty evaluation process that relates to their actual job duties, and connecting expectations to job descriptions would accomplish this goal.

In addition, the following policies should be reviewed at the system and college levels: 3.5.3 (Additional Activities), 3.5.4 (Professional Activities and Contributions), 3.6.0 (Teaching Effectiveness), and 3.6.1.3 (Criteria). VCCS policies should either (a) provide extensive lists of activities for each of these policies, with proposed levels of activity required for each faculty summary evaluation rating to provide a degree of uniformity across the colleges, or (b) the policies should be very general so the colleges can develop their own pertinent faculty activity lists that are negotiated and agreed upon by faculty members and their supervisors. As presently worded, the existing policy

statements provide no useful structure for the development of comprehensive faculty evaluation plans at the colleges. In addition, not all faculty members should be evaluated by the same list of activities.

Policy 3.6.1.5.0 requires annual evaluations of all faculty members after their first year of employment. It has been suggested that faculty working under multi-year contracts could be evaluated less frequently, either every other year or when they are under consideration for multi-year contract renewal. Such a policy change would lessen the time burden currently placed on faculty and administrators for everyone to engage in this process every year. This simple modification to existing policy has been discussed by the VCCS college vice presidents, but there has been no action to date to forward any suggested policy language changes for system-wide approval.

Arreola (2007) and Lee (2006) agree that the main problem with most performance evaluation plans is that they are aimed at rating performance (for promotion, tenure, etc.) versus appraising performance with the aim of working toward continuous improvement and professional development/enhancement. Logically, if a faculty member has taught successfully (i.e. has received meritorious evaluation ratings) for years, what do we think they will learn from another evaluation – if the rating has no real consequences? Faculty evaluation should be used for formative purposes with the goal of supporting continuous improvement through targeted professional development opportunities.

Currently, merit pay is regulated by the *VCCS Policy Manual* Section 3.8.11 (VCCS, 2010a), and a suggested policy change is to uncouple merit pay from the evaluation rating (whether the evaluation is conducted yearly or once every two, three, or

five years). True merit pay should be reserved “to recognize and reward past meritorious performance and to encourage future meritorious performance” (Arreola, 2007, p. 84). Awarding of merit pay could be a competitive process for which faculty members must complete an application, and merit pay could be awarded as a one-time bonus payment that would not be factored into a faculty member’s base salary. Merit pay applications should be reviewed by a small committee of faculty members and administrators at each campus, much like an ad hoc appointment advisory committee as defined in VCCS Policy 3.4.0.4 (VCCS, 2010a).

Merit pay applications would be conducted separately from the periodic formative faculty evaluation process. Under such a system, faculty would only be eligible to apply for merit pay raises during the years that they are up for a three- or five-year contract. Faculty members with meritorious evaluation rankings (good, very good, or excellent) who either do not apply for, or who do not receive approval for, merit pay would receive the base pay increase as approved by the state legislature and State Board for Community Colleges.

By uncoupling the formative rating from merit pay, the faculty evaluation process could become much more objective. The overall evaluation rating would still be utilized for summative promotion considerations, but the main purpose for conducting faculty evaluations would be for formative purposes and continuous improvement through professional development and growth activities.

College Review

Based on the results of this study, colleges should review their current faculty evaluation plans for alignment with the following guidelines. The plan should utilize

multiple measures of faculty performance, including contributions from student evaluations, supervisor evaluation, and reflective faculty narrative and/or portfolio self-evaluations. Other forms of input should be negotiated between individual faculty members and their supervisors. The VCCS and individual colleges need to develop objective, standardized criteria for measuring and reporting faculty member performance, with teaching receiving the greatest weighting in the evaluation plan. The criteria should be based on actual faculty teaching and service performance activities.

Professional Development Opportunities

There should be a connection between faculty evaluation results, professional development opportunities for continuous improvement/growth, and actual rewards/consequences that depend on the evaluation rating. The colleges and/or the VCCS should develop professional development training for evaluators on how to conduct a proper, objective evaluation and to inform supervisors about other professional development opportunities that are appropriate for faculty members.

Conclusions

While the results of this study supported the perception that teaching is the primary function of full-time community college teaching faculty members, the results also clearly showed that there is a gulf between faculty and administrator perceptions of the process, faculty activities and data sources that should be included in the faculty evaluation process, and the relative weighting or importance that should be assigned to each data element. Additional work needs to be done to bridge these differences and to build faculty evaluation processes that actually do what they are intended to do – provide

formative and summative information to faculty members and their supervisors (Arreola, 2007).

Faculty evaluation remains an antagonistic process for VCCS faculty members and administrators. As revealed through this study, many college plans have serious limitations, typically centered around four themes: (1) they are poorly designed and/or executed and not reflective of actual faculty performance activities, (2) they rely too heavily on one data source (usually student evaluations or supervisor evaluation), (3) the faculty evaluation process rarely distinguishes between truly outstanding faculty members and those who meet minimum expectations, and (4) there are usually no real consequences or rewards tied to the annual evaluation rating.

In addition to collecting, analyzing and reporting the results of this survey, one of the main goals of this study was to provide data to the colleges for use in the development/revision of their comprehensive faculty evaluation systems. It will be up to the individual college staff members whether or not they choose to utilize this data as they evaluate their existing plans for revision. On a more broad scale, the results of this study were to be used to create a “template” faculty evaluation plan for the VCCS that is built upon two foundation components: (a) best practices in the literature, and (b) actual data from VCCS teaching faculty members and their administrative supervisors. The best model in the literature is Arreola’s *Developing a Comprehensive Faculty Evaluation System: A Guide to Designing, Building, and Operating Large-Scale Faculty Evaluation Systems* (2007), and this study of VCCS faculty and administrator views on faculty evaluation supports much of the detailed material in Arreola’s book.

For faculty evaluation to be a meaningful exercise, VCCS policy should be supportive of good plan design and college plans should (a) be a reflection of actual faculty duties, (b) use multiple measures (including appropriate student data), (c) be based on objective, measurable criteria that distinguish between faculty members who meet minimum expectations and those who go above and beyond on a regular basis, and (d) the results of the objective analysis should provide (1) formative feedback to the faculty members related to their performance, (2) result in a professional development plan for continued improvement/growth, and 3) have tangible consequences tied to the overall rating. The process for awarding merit pay should be uncoupled from the faculty evaluation process and should be conducted as a separate process. By separating merit pay from the formative aspects of faculty evaluation, and tying performance ratings to job descriptions and objective criteria it is anticipated that faculty evaluation ratings will demonstrate more variance, thus allowing for true distinctions between different levels of faculty performance.

REFERENCES

- Addison, W. E., Best, J., & Warrington, J. D. (2006). Students' perceptions of course difficulty and their ratings of the instructor. *College Student Journal*, 40, 409-416.
- American Association of Community Colleges. (2009, December). *Fast facts*. Retrieved June 2, 2010, from the AACC Web site
<http://www.aacc.nche.edu/AboutCC/Pages/fastfacts.aspx>
- American Association of University Professors. (2006). *Statement on teaching evaluation*. Retrieved May 13, 2008, from the AAUP Web site
<http://www.aaup.org/NR/rdonlyres/2BE9B083-8942-4BB9-8315-8BF772CB57F9/0/StatementonTeachingEvaluation.pdf>
- Appling, S. E., Naumann, P. L., & Berk, R. A. (2001). Using a faculty evaluation triad to achieve evidence-based teaching. *Nursing and Health Care Perspectives*, 22, 247-251.
- Arreola, R. A. (2005, April). *Validity, like beauty is...* Paper presented at the meeting of the 2005 American Educational Research Association Convention, Montreal.
- Arreola, R. A. (2007). *Developing a comprehensive faculty evaluation system*. Boston, MA: Anker.
- Baldwin, T., & Blattner, N. (2003). Guarding against potential bias in student evaluations: What every faculty member needs to know. *College Teaching*, 51, 27-32.
- Basow, S. A., & Silberg, N. T. (1987). Student evaluations of college professors: Are female and male professors rated differently? *Journal of Educational Psychology*, 79, 308-314.

- Berk, R. A. (2006). *Thirteen strategies to measure college teaching*. Sterling, VA: Stylus.
- Birnbaum, M. H. (1999). A survey of faculty opinions concerning student evaluations of teaching. Retrieved April 18, 2004, from the University of California, Fullerton Web site <http://psych.fullerton.edu/mbirnbaum/faculty3.htm>
- Blum, M. L. (1936). An investigation of the relation existing between students' grades and their ratings of the instructors' ability to teach. *Journal of Educational Psychology*, 27, 217-222.
- Brock, C. M, Chrestman, C., & Armstrong, J. (1998). *Itawamba Community College presents a system of administrative evaluation for faculty retention: A systematic approach to faculty evaluation and retention* (Report No. JC990477). Itawamba Community College. (ERIC Document Reproduction Service No. ED432344)
- Campion, W. J., Mason, D. V., & Erdman, H. (2000). How faculty evaluations are used in Texas community colleges. *Community College Journal of Research and Practice*, 24, 169-179.
- Cashin, W. E. (1999). Student ratings of teaching: Uses and Misuses. In P. Seldin & Associates *Changing practices in evaluating teaching: A practical guide to improved faculty performance and promotion/tenure decisions* (pp. 25-44). Bolton, MA: Anker.
- Cashin, W. E., & Downey, R. G. (1992). Using global student rating items for summative evaluation. *Journal of Educational Psychology*, 84(4), 563-572.
- Centra, J. A. (n.d.) *Will teachers receive higher student evaluations by giving higher grades and less coursework?* Retrieved April 19, 2004 from the Educational Testing Service Web site <http://www.ets.org/Media/Products/StudentEval.pdf>

- Centra, J. A. (1993). *Reflective faculty evaluation: enhancing teaching and determining faculty effectiveness*. San Francisco, CA: Josey-Bass.
- Chickering, A. W., & Gamson, Z. F. (1987, March). Seven principles for good practice in undergraduate education. *AAHE Bulletin*, 39(7), 3-7.
- Cutler, A. (2007). A night at the circus. *Journal of College Science Teaching*, 36(7), 6-7.
- Davidovich, N., & Soen, D. (2006). Class Attendance and students' evaluation of their college instructors. *College Student Journal*, 40, 691-703.
- Dee, K. C. (2007). Student perceptions of high course workloads are not associated with poor student evaluations of instructor performance. *Journal of Engineering Education*, 96, 69-78.
- Donnelly, J. R. (2006). Student ratings of faculty: Differences by employment status and grades given. *Dissertation Abstracts International*, 67 (06), 2068A. (UMI No. 3220509)
- Downey, J. A. (2003). Emotional awareness as a mediator of community college student satisfaction ratings. *Community College Journal of Research and Practice*, 27, 711-720.
- Educational Testing Service (2006a). *Student instructional report II (SIR II): Its development, uses and supporting research (2006)*. Retrieved April 19, 2008, from the Educational Testing Service Web site:
http://www.ets.org/Media/Products/SIR_II/pdf/53228_sirII_white_paper.pdf
- Educational Testing Service (2006b). *Student instructional report II (SIR II): Comparative data (2006)*. Retrieved April 19, 2008, from the Educational Testing Service Web site <http://www.ets.org/Media/Products/Sircomp2.pdf>

- Fairweather, J. S. (2002). In C. L. Colbeck (Ed.) *Evaluating Faculty Performance* (pp. 97-108). San Francisco, CA: Jossey-Bass.
- Fenwick, T. J. (2001). Using student outcomes to evaluate teaching: A cautious exploration. *New Directions for Teaching and Learning*, 88, 63-74.
- Fitzpatrick, J. L., Sanders, J. R., & Worthen, B. R. (2004). *Program valuation: Alternative approaches and practical guidelines* (3rd ed.). Boston: Pearson.
- Fleak, S. K., Romine, J., & Gilchrist, N. (2003). Portfolio peer review: A tool for program change. *Journal of Education for Business*, 72, 139-146.
- Foote, D. A., Harmon, S. K., & Mayo, D. T. (2003). The impacts of instructional style and gender role attitude on students' evaluation of faculty. *Marketing Education Review*, 13, 9-19.
- Germain, M., & Scandura, T. A. (2005). Grade inflation and student individual differences a systematic bias in faculty evaluations. *Journal of Instructional Psychology*, 32, 58-67.
- Glenn, D. (2007, May 29). Method of using student evaluations to assess professors is flawed but fixable, 2 scholars say. *The Chronicle of Higher Education*, p. A29.
- Grubb, W. N., Worthen, H., Byrd, B., Webb, E., Badway, N., Case, C., et. al. (1999). *Honored but invisible: An inside look at teaching in community colleges*.
- Harbour, C. P., & Jaquette, O. (2007). Advancing an equity agenda at the community college in an age of privatization, performance accountability, and marketization. *Equity & Excellence in Education*, 40, 197-207.
- Hobson, S. M., & Talbot, D. M. (2001). Understanding student evaluations. *College Teaching Journal*, 49. Retrieved April 28, 2004, from <http://80->

infotrac.galegroup.com.proxy.lib.odu.edu/itw/infomark/498/745/51086878w4/purl=rc1_EAIM_0_A71188224&dyn=6!xrn_1_0_A71188224?sw_aep=viva_odu

Jaquette, O. (2006, February). Making performance accountability work: English lessons for U.S. community colleges. Retrieved June 20, 2010, from <http://www.jff.org/publications/education/making-performance-accountability-work-e/345>

Jenkins, J. D. (2006). Toward the formulation of a community college policy for the evaluation of teaching based on evidence of student learning. *Dissertation Abstracts International*, 67 (06), 2026A. (UMI No. 3220794)

Koops, J. B., & Winsor, K. A. (2006). Creating a professional learning culture through faculty evaluation. *The Journal of Education*, 186, 61 -70.

Langbein, L. I. (1994). The validity of student evaluations of teaching. *P.S.: Political Science and Politics*, 27, 545-553.

Lee, C. D. (2006). *Performance conversations: An alternative to appraisals*. Tucson, AZ: Fenestra.

Maricopa Community College District (2007). *Faculty evaluation plan*. Retrieved March 13, 2007, from the Maricopa Community College District Web site <http://www.maricopa.edu/org/faculty/FEP/FEP%20Text.htm>.

Maricopa Community College District (2009). *Summary of fiscal year headcount*. Retrieved March 3, 2009, from the Maricopa Community College District Web site http://www.maricopa.edu/business/ir/trends/HTML/FY_DO.htm.

- McArthur, R. C. (2002). Democratic leadership and faculty empowerment at the community college: A theoretical model for the department chair. *Community College Review*, 30(3), 1-10.
- McCaig, R. D. (2002). Publicize or perish? College faculty perspectives on the validity of and public access to student ratings of teaching effectiveness. *Dissertation Abstracts International*, 63 (06), 2068A. (UMI No. 3057093)
- McGee, R. L. (1996). Factors in faculty evaluation procedures in community colleges. *Dissertation Abstracts International*, 57 (07), 2824A. (UMI No. 9636882)
- McLean, C. P. (1994). Community college faculty performance appraisal: A multi-perception analysis. *Dissertation Abstracts International*, 55 (11), 3393A. (UMI No.9510411)
- McPherson, M. A. (2006). Determinants of how students evaluate teachers. *Journal of Economic Education*, 37, 3-20.
- Melland, H. I., & Volden, C. M. (1996). Teaching portfolios for faculty evaluation. *Nurse Educator*, 21(20), 35-38.
- Millea, M., & Grimes, P. W. (2002). Grade expectations of student evaluation of teaching. *College Student Journal*, 36, 582-590.
- Milton, O., & Shoben, E. J. (Eds.) (1968). *Learning and the professors*. Athens, OH: Athens University Press.
- Morris, G. T. (1997). The design and implementation of a faculty evaluation system. *Dissertation Abstracts International*, 58, (12), 4582A. (UMI No. 9817888)
- Murray, J. P. (2002). Faculty development in SACS-accredited community colleges. *Community College Review*, 29(4), 50-66.

- Myers, C. B. (2008). College faculty and the scholarship of teaching: Gender differences across four key activities. *Journal of the Scholarship of Teaching and Learning*, 8(2), 38-51.
- Nuhfer, E. (2005, April). *A fractal thinker looks at student evaluations*. Paper presented at the meeting of the 2005 American Educational Research Association Convention, Montreal.
- Ory, J. C. (2001). Faculty thoughts and concerns about student ratings. In K. Lewis (Ed.) *Techniques and strategies for interpreting student evaluations* (pp. 3-16). San Francisco: Jossey-Bass.
- Osborne, J. L. (1998). Integrating student and peer evaluation of teaching. *College Teaching*, (46), 36-38.
- Paulsen, M. B. (2002). Evaluating teaching performance. In C. L. Colbeck (Ed.) *Evaluating Faculty Performance* (pp. 5-18). San Francisco, CA: Jossey-Bass.
- Read, W. J., Rama, D. V., & Raghunandan, K. (2001). The relationship between student evaluations of teaching and faculty evaluations. *Journal of Education for Business*, 76(4). Retrieved April 24, 2004, from http://80-infotrac.galegroup.com.proxy.lib.odu.edu/itw/infomark/498/745/51086878w4/purl=rc1_EAIM_0_A76133441&dyn=15!xrn_1_0_A76133441?sw_aep=viva_odu
- Redmon, K. (1999). ERIC review - Faculty evaluation: A response to competing values. *Community College Review*, 27, 57-71.
- Remmers, H. H. (1930). To what extent do grades influence student ratings of instruction? *Journal of Educational Research*, 21, 314-316.

- Russell, A., & Gadberry, L. (2000). *Student opinion of faculty evaluations spring 2000 and student opinion of faculty evaluations a five year comparison: Fall 1994 and spring 2000*. (Report No. HE033561). Southwestern Oklahoma State University. (ERIC Document Reproduction Service No. ED446706)
- Sain, B.P. (2008). *Perceptions of portfolios for faculty evaluation in North Carolina community colleges*. Unpublished doctoral dissertation, Western Carolina University, Cullowhee, NC. (UMI No. 3329385)
- Schaffner, M., & MacKinnon, F. J. D. (2002). *A standards-driven approach to faculty evaluation: The conflict of change*. (Report No. HE035148). Bowling Green State University. (ERIC Document Reproduction Service No. ED467600)
- Scriven, M. (2005, April). *Critical issues in faculty evaluation: Valid data and the validity of practice*. Paper presented at the meeting of the 2005 American Educational Research Association Convention, Montreal.
- Seldin, P. (1984). *Changing practices in faculty evaluation*. San Francisco: Jossey-Bass.
- Seldin, P. (1993, October). Rating student ratings of professors. *Education Digest*, 59(2), 70. Retrieved May 14, 2008, from <http://search.ebscohost.com.proxy.lib.odu.edu/login.aspx?direct=true&db=ehh&AN=9309287677&site=ehost-live>
- Seldin, P., & Associates. (1999). *Changing practices in evaluating teaching: A practical guide to improved faculty performance and promotion/tenure decisions*. Bolton, MA: Anker.

Sojka, J., Gupta, A. K., & Deeter-Schmetz, D. R. (2002). Student and faculty perceptions of student evaluations of teaching, a study of similarities and differences. *College Teaching*, 50, 44-49.

Southern Association of Colleges and Schools, Commission on Colleges (1997, December). *Criteria for accreditation: Commission on colleges*. Retrieved March 23, 2009 from <http://wise.fau.edu/sacs/crit.pdf>

Southern Association of Colleges and Schools, Commission on Colleges (2009, December). *Principles of accreditation: Foundations for quality enhancement*. Retrieved June, 2, 2010, from <http://www.sacscoc.org/pdf/2010principlesofacreditation.pdf>

South Texas College faculty evaluation plan. (2004) Retrieved January 26, 2009, from http://academicaffairs.southtexascollege.edu/facresources/evalplan/pdf/fac_eval_plan_10_06.pdf

State Council of Higher Education for Virginia (SCHEV). (2010). *Public institution enrollment totals*. Retrieved June 2, 2010, from http://research.schev.edu/enrollment/public_totals.asp

Stemler, S. (2001). An overview of content analysis, *Practical Assessment, Research & Evaluation* 7(17). Retrieved August 7, 2009, from <http://PAREonline.net/getvn.asp?v=7&n=17>

Stryker, S., & Serpe, R. T. (1982). Commitment, identity salience, and role behavior: A theory and research example. In W. Ickes and E. S. Knowles (Eds.), *Personality, Roles, and Social Behavior* (pp. 199-218). New York: Springer-Verlag.

Summers, M. (2007, November). *Chancellor's faculty evaluation task force: A summary report*. Presented at the meeting of the VCCS Academic and Student Services Council, Roanoke, VA.

Theall, M. (2005, April). *Valid faculty evaluation data: Are there any?* Paper presented at the meeting of the 2005 American Educational Research Association Convention, Montreal.

Virginia Community College System (2003). *A strategic direction: Dateline 2009*.

Retrieved October 2, 2007, from

<http://www.vccs.edu/ChancellorsDatelineVision/tabid/426/Default.aspx>

Virginia Community College System (2008a). *Chancellor's advisory council of presidents meeting minutes, April 2008*. Retrieved September 22, 2008 from the Virginia Community College System Web site

<http://www.vccs.edu/Portals/0/ContentAreas/Meetings/ACOP-Minutes-April2008.pdf>

Virginia Community College System (2008b). *Annual enrollments by college, 2007-2008*. Retrieved February 8, 2009 from the Virginia Community College System Web site <http://myfuture.vccs.edu/Research/an0708.htm>

Virginia Community College System (2008c). *Salaries, tenure, and fringe benefits of full-time instructional faculty: Academic year 2007-2008*. Retrieved February 8, 2009 from the Virginia Community College System Web site

http://myfuture.vccs.edu/Research/IPEDS_SA_FA07.pdf

Virginia Community College System (2009). *State Board for Community College Minutes*. Retrieved March 8, 2009 from the Virginia Community College System

Web site

<http://www.vccs.edu/WhoWeAre/StateBoard/Minutes/tabid/144/Default.aspx>

Virginia Community College System (2010a). *Virginia Community College System Policy Manual, Section 3*. Retrieved June 2, 2010 from the Virginia Community College System Web site

<http://www.vccs.edu/Portals/0/ContentAreas/PolicyManual/Sct3.pdf>

Virginia Community College System (2010b). Virginia's Community Colleges Fast Facts. Retrieved June 2, 2010 from the Virginia Community College System Web site <http://www.vccs.edu/WhoWeAre/FastFacts/tabid/79/Default.aspx>

Virginia Community College System (2010c). *State Board for Community College Minutes*. Retrieved June 2, 2010 from the Virginia Community College System Web site <http://www.vccs.edu/Portals/0/ContentAreas/StateBoard/Mar2010SBFinalMin.pdf>

Wolfer, T. A., & Johnson, M. M. (2003). Re-evaluating student evaluation of teaching: The teaching evaluation form. *Journal of Social Work Education, 39*, 111-121.

Worcester, P. I. (1993). Faculty and administrator perceptions of faculty evaluation in a technical college. *Dissertation Abstracts International, 54* (03), 789A. (UMI No. 9319419)

Wright, R. E. (2006). Student evaluations of faculty: Concerns raised in the literature and possible solutions. *College Student Journal, 40*, 417-422.

- Yao, Y., & Grady, M. L. (2006). How do faculty make formative use of student evaluation feedback?: A multiple case study. *Journal of Personnel Evaluation in Education, 18*, 107-126.
- Yon, M., Burnap, C., & Kohut, G. (2002). Evidence of effective teaching: perceptions of peer reviewers. *College Teaching, 50*, 104-111.
- Yunker, P. J., & Yunker, J. A. (2003). Are student evaluations of teaching valid? Evidence from an analytical business core course. *Journal of Education for Business, 78*, 313-318.

Appendix A
Mail Correspondence to College Presidents

RE: Request to contact your academic deans and teaching
Faculty members for participation in my doctoral study

Dear Dr. _____ -

Please accept this formal request for your permission to survey the academic deans and teaching faculty members at your college on the topic of faculty evaluation. This survey is the subject of my dissertation research and is also in alignment with the ASAC, CFAC and Council of Deans interests in improving the VCCS faculty evaluation process. I am a doctoral candidate in the Community College Leadership program at Old Dominion University and the title of my dissertation is: *Perceptions of Virginia Community College System Faculty and Administrators on Faculty Evaluation*.

The purpose of my study is to fill significant voids in the literature by providing information on: (1) community college teaching faculty and administrator perceptions on the faculty evaluation process, and (2) how those perceptions differ based on role, gender, content area and years of experience. The results of my study may also be useful the next time your college reviews your current faculty evaluation plan.

The participation and support of your college staff will be most appreciated as I complete the final steps in earning my Ph.D. You have my assurance that: (1) the time to complete the survey will be kept to a minimum – approximately 15 minutes, (2) survey participants will remain completely anonymous, (3) only aggregate data will be reported in my dissertation, and (4) aggregate results will be shared with you. My study has already received Human Subjects approval from Old Dominion University and the VCCS Office of Institutional Effectiveness.

May I have your permission to contact and survey the academic deans and full-time teaching faculty members at your college? Enclosed is a postage-paid card for your convenience in responding to this request – or you may send an email response if you prefer to: bhightower@vccs.edu. If you have further questions, please contact me or my dissertation chair, Dr. Linda Bol, at the contact information listed below. Thank you for your assistance.

Respectfully,

William H. Hightower, Jr.
Director of Educational Programs
Virginia Community System Office
ODU Doctoral Candidate
(804) 819-4696 work
(804) 840-7565 cell
bhightower@vccs.edu

Dr. Linda Bol
Professor of Educational Foundations
Old Dominion University
Darden College of Education
Faculty Advisor
Office number: (757) 683-4584
lbol@odu.edu

Appendix B
Email Correspondence to College Academic Deans

Dear _____:

Recently, I received permission from ____ to include your college in the research I am conducting toward my dissertation. I am a doctoral candidate in the Community College Leadership program at Old Dominion University and the title of my dissertation is: *Perceptions of Virginia Community College System Faculty and Administrators on Faculty Evaluation*.

The purpose of my study is to fill significant voids in the literature by providing information on: (1) community college faculty and administrator perceptions on the faculty evaluation process, and (2) how those perceptions differ based on role, gender, content area and years of experience. The results of my study may also be useful when your college staff reviews your current faculty evaluation plan. My research includes surveying all academic deans and full-time teaching faculty members at all participating VCCS colleges through the use of an online survey instrument.

Specifically, I am asking you to complete the on-line survey instrument by clicking on the following link: <https://survey.vccs.edu/ss/wsb.dll/s/41g732>. In addition, please forward this message to all of the full-time teaching faculty members whom you supervise to request that they also complete the survey. Participation in this study is voluntary - participants may refuse to answer any questions they do not want to answer (after the first two questions) and still remain in the study. Participants may withdraw from the study at any time without consequences.

Your participation and support will be most appreciated as I complete the final steps in earning my Ph.D. You have my assurance that: (1) the time to complete the survey will be kept to a minimum – approximately 15 minutes, (2) survey participants will remain anonymous, (3) only aggregate data will be reported in my dissertation, and (4) aggregate results will be provided to your president. If the number of participants at any one college is so small that demographic information would lead to the identification of specific individuals, only system-wide data will be shared with your president. In addition, summary data will be provided at future CODD and CFAC meetings. My study has received Human Subjects Committee approval from ODU. For interested survey participants, there will be an opportunity to enter into a random drawing for one of several Barnes & Noble gift cards. Your participation is crucial to this study and is greatly appreciated. If you have any questions, please contact either myself or my dissertation advisor as indicated below. Thank you for your time.

Respectfully,

William H. Hightower, Jr.
Director of Educational Programs
Virginia Community System Office
ODU Doctoral Candidate
(804) 819-4696 work
bhightower@vccs.edu

Dr. Linda Bol
Professor of Educational Foundations
Old Dominion University
Darden College of Education
Office number: (757) 683-4584
lbol@odu.edu

Appendix C
On-line Faculty Evaluation Survey

VCCS Faculty Evaluation Survey

Thank you for visiting the VCCS Faculty Evaluation Survey. Participation in this study is voluntary and completely anonymous - participants may refuse to answer any questions they do not want to answer and still remain in the study. Participants may withdraw from the study at any time without consequences.

Your participation and support will be most appreciated as I complete the final steps in earning my Ph.D. You have my assurance that: (1) the time to complete the survey will be kept to a minimum – approximately 15 minutes, (2) survey participants will remain anonymous, (3) only aggregate data will be reported in my dissertation, and (4) your aggregate college-specific results will be provided to your president. If the number of participants at any one college is so small that demographic information would lead to the identification of specific individuals, only system-wide data will be shared with your president. In addition, summary data will be provided at future CODD and CFAC meetings.

Please select your college from the dropdown list.

What is your **PRIMARY** role?

- ☐ Teaching faculty ☐ Administrator

[NOTE: Respondents who select “Teaching faculty” will be redirected to the following page:]

What is your **PRIMARY** area of responsibility?

- ☐ Developmental education
- ☐ Career and technical education
- ☐ Transfer/general education
- ☐ Other (please specify)

If you selected other, please specify:

For how many years have you been a full-time teaching faculty member at the post-secondary level?

- ☐ 0
- ☐ 1 - 3
- ☐ 4 - 6
- ☐ 7 - 9
- ☐ 10 - 12
- ☐ 13 - 15
- ☐ 15 +

For how many years have you been a full-time teaching faculty member at your current institution?

- ☐ 0
- ☐ 1 - 3
- ☐ 4 - 6
- ☐ 7 - 9
- ☐ 10 - 12
- ☐ 13 - 15
- ☐ 15 +

What is your gender?

- ☐ Female
- ☐ Male

[NOTE: Respondents who select “Administrator” will be redirected to the following page:]

What is your **PRIMARY** area of responsibility?

- ☐ Developmental education
- ☐ Career and technical education
- ☐ Transfer/general education
- ☐ All instructional areas

For how many years have you been responsible for evaluating teaching faculty members at the post-secondary level?

- ☐ 0
- ☐ 1 - 3
- ☐ 4 - 6
- ☐ 7 - 9
- ☐ 10 - 12
- ☐ 13 - 15
- ☐ 15 +

For how many years have you been responsible for evaluating teaching faculty members at your current institution?

- ☐ 0
- ☐ 1 - 3
- ☐ 4 - 6
- ☐ 7 - 9
- ☐ 10 - 12
- ☐ 13 - 15
- ☐ 15 +

What is your gender?

- ☐ Female
- ☐ Male

[NOTE: From this point forward, all respondents will be presented with the following questions:]

PURPOSE AND USE OF AN IDEAL FACULTY EVALUATION PLAN

DIRECTIONS: Please indicate your agreement with each of the following statements related to the ***purpose(s) and use of an ideal faculty evaluation plan*** by choosing ONE response for each ITEM.

Faculty evaluation should be primarily a **FORMATIVE** process (one that is designed to enhance continuous improvement in performance of teaching faculty members).

☐ strongly disagree ☐ disagree ☐ agree ☐ strongly agree

Faculty evaluation should be primarily a **SUMMATIVE** process (one that is designed to provide an overall performance rating that will be used by administrators as the basis for making personnel decisions such as determining contract eligibility, promotion, and merit pay status).

☐ strongly disagree ☐ disagree ☐ agree ☐ strongly agree

Faculty evaluation should be **BOTH formative and summative**, serving to enhance continuous improvement in teaching faculty performance while also serving as the basis for personnel decisions.

☐ strongly disagree ☐ disagree ☐ agree ☐ strongly agree

Faculty evaluation should be tied to an individual professional development plan to focus efforts on any areas of the faculty member's performance that could use improvement.

☐ strongly disagree ☐ disagree ☐ agree ☐ strongly agree

Evaluators of teaching faculty members should have some degree of discretion/judgment with regard to the assignment of annual summative faculty evaluation ratings.

☐ strongly disagree ☐ disagree ☐ agree ☐ strongly agree

Two or more successive faculty evaluation ratings of "excellent" should be a requirement for promotion to associate professor and full professor.

☐ strongly disagree ☐ disagree ☐ agree ☐ strongly agree

Merit pay should be directly tied to the annual summative faculty evaluation rating.

☐ strongly disagree ☐ disagree ☐ agree ☐ strongly agree

FORMATIVE EVALUATION OF OVERALL FACULTY PERFORMANCE

DIRECTIONS: To indicate the major evaluative elements that should be utilized in evaluating faculty members *for the purpose of improving their teaching performance*, please indicate the *relative importance* of including each of the following factors in your college faculty evaluation plan by choosing ONE response for each of these major evaluative elements.

Teaching.

☐ not important ☐ somewhat important ☐ important ☐ very important

Service to the college, discipline or community.

☐ not important ☐ somewhat important ☐ important ☐ very important

Scholarly or creative activities (such as publications, performances or works of art).

☐ not important ☐ somewhat important ☐ important ☐ very important

Professional recognition or accomplishment.

☐ not important ☐ somewhat important ☐ important ☐ very important

Personal attributes (such as collegiality, interpersonal relationships, punctuality, ability to follow instructions, etc.).

☐ not important ☐ somewhat important ☐ important ☐ very important

Other elements (please specify).

SUMMATIVE EVALUATION OF OVERALL FACULTY PERFORMANCE

DIRECTIONS: What major evaluative elements should be utilized in evaluating faculty members *for promotion in rank, salary increase, or contract considerations?* Please indicate the *relative importance* of each factor by choosing ONE response for each item number.

NOTE: Responses in this section may differ from the answers provided in the previous section due to the different purposes for conducting formative and summative faculty evaluations.

Teaching.

☐ not important ☐ somewhat important ☐ important ☐ very important

Service to the college, discipline or community.

☐ not important ☐ somewhat important ☐ important ☐ very important

Scholarly or creative activities (such as publications, performances or works of art).

☐ not important ☐ somewhat important ☐ important ☐ very important

Professional recognition or accomplishment.

☐ not important ☐ somewhat important ☐ important ☐ very important

Personal attributes (such as collegiality, interpersonal relationships, punctuality, ability to follow instructions, etc.).

☐ not important ☐ somewhat important ☐ important ☐ very important

Other elements (please specify).

EVALUATION OF TEACHING PERFORMANCE

DIRECTIONS: Please indicate the importance of each factor for evaluating a faculty member's **TEACHING** performance by choosing ONE response for each item.

Classroom Performance

Enrollment in elective courses.

☐ not important ☐ somewhat important ☐ important ☐ very important

Delivering lectures.

☐ not important ☐ somewhat important ☐ important ☐ very important

Supervising laboratory sessions.

☐ not important ☐ somewhat important ☐ important ☐ very important

Operating a chat room or discussion board.

☐ not important ☐ somewhat important ☐ important ☐ very important

Facilitating small group experiential learning events.

☐ not important ☐ somewhat important ☐ important ☐ very important

Preparation for Class/Laboratory/Clinical/etc.

Course syllabi and examinations.

☐ not important ☐ somewhat important ☐ important ☐ very important

Developing course materials (e.g. handouts, lab exercises, computer simulations, experiential learning events, etc.).

☐ not important ☐ somewhat important ☐ important ☐ very important

Creating an on-line course.

☐ not important ☐ somewhat important ☐ important ☐ very important

Developing written examinations.

☐ not important ☐ somewhat important ☐ important ☐ very important

Grading examinations.

☐ not important ☐ somewhat important ☐ important ☐ very important

Course load.

☐ not important ☐ somewhat important ☐ important ☐ very important

Student Performance & Evaluation of Faculty

Student examination performance.

☐ not important ☐ somewhat important ☐ important ☐ very important

Grade distributions.

☐ not important ☐ somewhat important ☐ important ☐ very important

Systematic student evaluations.

☐ not important ☐ somewhat important ☐ important ☐ very important

Informal student opinions.

☐ not important ☐ somewhat important ☐ important ☐ very important

Self Evaluation

Self-evaluation checklist or rating scale.

☐ not important ☐ somewhat important ☐ important ☐ very important

Self-evaluative narrative report.

☐ not important ☐ somewhat important ☐ important ☐ very important

Compilation of detailed portfolio of faculty and student accomplishments.

☐ not important ☐ somewhat important ☐ important ☐ very important

Additional/External Evaluation

Peer evaluation.

☐ not important ☐ somewhat important ☐ important ☐ very important

Direct supervisor's evaluation.

☐ not important ☐ somewhat important ☐ important ☐ very important

Dean evaluation.

☐ not important ☐ somewhat important ☐ important ☐ very important

VP/Provost evaluation.

☐ not important ☐ somewhat important ☐ important ☐ very important

Alumni evaluation.

☐ not important ☐ somewhat important ☐ important ☐ very important

Committee evaluation.

☐ not important ☐ somewhat important ☐ important ☐ very important

Other elements related to teaching performance (please specify).

EVALUATION OF SERVICE PERFORMANCE

DIRECTIONS: Please indicate the importance of each factor for evaluating a faculty member's *service performance* by choosing ONE response for each item.

Service to Students

Maintaining regular office hours.

☐ not important ☐ somewhat important ☐ important ☐ very important

Academic advising.

☐ not important ☐ somewhat important ☐ important ☐ very important

Non-academic student counseling.

☐ not important ☐ somewhat important ☐ important ☐ very important

Willingness to teach "undesirable" courses.

☐ not important ☐ somewhat important ☐ important ☐ very important

Advisor to student organizations.

☐ not important ☐ somewhat important ☐ important ☐ very important

Participation in campus programs.

☐ not important ☐ somewhat important ☐ important ☐ very important

Service to the College

Service on department committees.

☐ not important ☐ somewhat important ☐ important ☐ very important

Service on college committees.

☐ not important ☐ somewhat important ☐ important ☐ very important

Departmental administrative duties.

☐ not important ☐ somewhat important ☐ important ☐ very important

Service as student recruiter.

☐ not important ☐ somewhat important ☐ important ☐ very important

Service to the Community

Service on local community boards.

☐ not important ☐ somewhat important ☐ important ☐ very important

Participation in local activities/events.

☐ not important ☐ somewhat important ☐ important ☐ very important

Provision of discipline-related expertise to community groups (e.g. unpaid speaking engagements or consulting).

☐ not important ☐ somewhat important ☐ important ☐ very important

Service as a judge for local civic or school events.

☐ not important ☐ somewhat important ☐ important ☐ very important

Other types of service performance (please specify).

SUMMARY COMMENTS ON THE FACULTY EVALUATION PROCESS

DIRECTIONS: Please respond to the following open-ended questions/statements using the provided textboxes.

What are the strengths of the faculty evaluation process at your college?

What are the limitations of the faculty evaluation process at your college?

What needs to be changed with the faculty evaluation process at your college?

Please use the following space to address any issues related to faculty evaluation that you feel were not adequately covered in this survey instrument.



Thank you for participating in my study. To complete the survey and to learn about an opportunity to win a \$25 Barnes & Noble gift card, click the "submit" button, below.

Submit Survey  100%

Appendix D

GIFT CARD LOTTERY FOR SURVEY PARTICIPANTS

Barnes & Noble Gift Card Lottery

Thank you for completing the VCCS Faculty Evaluation Survey! If you would like to be entered into a drawing for one of several \$25 Barnes & Noble Gift Cards, please complete the following entry form. This drawing has no connection to your survey responses - the survey is completely anonymous and no attempt will be made to identify any survey participant nor to connect gift card lottery participants to any response submitted on the survey.

1) Please provide your name in the following box (Last name, First name).

2) Please provide your college name in the following box.

3) Please provide an email address where I may contact you over the summer if you are a gift card winner.

Odds of winning a gift card are dependent on the number of lottery entrants. Good luck with the drawing and, again, thank you for responding to the survey questions and providing data for my doctoral dissertation.

Respectfully, Bill Hightower.

[Submit Survey](#)

Appendix E

VCCS POLICY MANUAL SECTIONS 3.5 AND 3.6

The purpose of this section is to record the various personnel rules, regulations, policies, and procedures of the Federal and State governments, the State Board, and the System Office. Special attention is given to the difference in provisions for faculty and classified employees.

3.5 Faculty Responsibilities (C)

The major emphasis shall be on teaching, by working with students in classrooms, laboratories, individual conferences, and related activities to help the students develop their interests and abilities to the fullest capacity to become better persons, better workers, and better citizens. To accomplish this goal, the following work loads are expected of faculty.

3.5.0 Classes (C)

Faculty teaching loads during the academic year shall include such combinations of day, evening, and weekend classes as the needs of the college require. Twelve-fifteen (12-15) credit hours and fifteen-twenty (15-20) contact hours per semester are required for all full-time faculty. For the purpose of workload calculations, every lecture hour shall equate to one (1) credit hour and one (1) contact hour; and every laboratory hour shall equate to one-half (1/2) credit hour and one (1) contact hour. When the number of credit hours falls below twelve (12) because of the number of laboratory hours involved, the number of contact hours should be increased to bring the teaching load to the minimum of twelve (12) credit hours (utilizing the standard of two (2) laboratory hours equal one (1) credit hour) or to a maximum of twenty-four (24) contact hours.

Faculty teaching loads shall be calculated for the academic year, with a teaching load less than or in excess of normal for the fall semester being compensated for with adjustments in teaching load in the spring semester.

A faculty teaching load may also be adjusted by the college to take into consideration such factors as the use of instructional assistance, team teaching, the use of non-traditional instructional delivery systems, special assignments, and curriculum development. Curriculum development should be primarily for the development of a new program or new course in a program and/or the complete revision of an existing course or program.

Teaching-load adjustments shall be expressed in terms of an equivalent teaching load for the purpose of computing a faculty member's total teaching load.

3.5.1 Office Hours (C)

To promote the availability of faculty to work with individual students, all full-time faculty members are required to post on or near their office doors a minimum of 10 hours per week as office hours to be available to work with students on their individual academic and occupational problems. All adjunct faculty are required to provide for student advising and related activities. Each adjunct faculty member shall ensure that all students have been informed of the contact details including location and time.

Exception due to distance learning, off-campus assignments, or use of technology to serve students may be approved by the Academic Vice President or designee.

3.5.2 Teaching Faculty Assigned Temporary Administrative/ Professional Duties

Regular nine and twelve month teaching faculty may be temporarily assigned non-teaching duties (released time) for administrative/professional activities of more than 50% of an individual's full-time teaching load for a maximum of two academic years by the college president. Faculty assigned more than 50% released time for non-teaching duties for more than two years must be classified as administrative faculty unless an extension beyond two years is approved by the Chancellor. The college shall maintain a record of all released time for audit purposes.

3.5.3 Additional Activities (C)

Faculty responsibilities include committee work, student activities, community activities, student advising, and professional activities.

3.5.4 Professional Activities and Contributions

In addition to teaching effectiveness, faculty are expected to engage in and contribute toward the good of the college and its community. This requires that faculty members maintain current competence in their disciplines or specializations and that they share their expertise, time, and talents with the larger college community. Performance in this category will be measured not only by membership or affiliation but also by the quality of the contributions made by faculty members toward these endeavors. Such activities may include but are not limited to:

- a. Membership and activity in professional and civic organizations (general and/or specialized organizations at the local, state, and/or national levels);
- b. The accomplishment of important professional development activities that may or may not be part of an individual professional development plan;
- c. Attending and participating in professional conferences; workshops, and meetings;

- d. Keeping current regarding developments in education and industry;
- e. Participating in business or industrial activities related to professional field;
- f. Participating in college and state-level professional development activities;
- g. Being active in college and Systemwide committees;
- h. Engaging in writing speeches and reports and in consulting;
- i. Engaging in classroom-based research to improve teaching or in discipline-based research that may lead to publication;
- j. Sharing innovations in using instructional technology with colleagues in other colleges;
- k. Participating in the community service program at the colleges;
- l. Participating in local colleges advisory committees; and
- m. Contributing to community welfare and community development.

3.5.5 Academic Freedom and Responsibility (SB)

To ensure the college an instructional program marked by excellence, the Virginia Community College System supports the concept of academic freedom. In the development of knowledge, research endeavors, and creative activities, college faculty and students must be free to cultivate a spirit of inquiry and scholarly criticism.

Faculty members are entitled to freedom in the classroom in discussing their subjects, but should be careful not to introduce teaching matters which have no relation to their fields. Faculty and students must be able to examine ideas in an atmosphere of freedom and confidence and to participate as responsible citizens in community affairs.

The System also recognizes that commitment to every freedom carries with it attendant responsibilities. Faculty members must fulfill responsibility to society and to their profession by manifesting academic competence, professional discretion, and good citizenship. When they speak or write as a citizen, they will be free from institutional censorship or discipline, but their special position in the community imposes special obligations. As professional educators, they must remember that the public may judge their profession and their institution by their utterances. Hence, they should at all times be accurate, exercise appropriate restraint, show respect for the opinions of others, and make every effort to indicate that they are not an institutional spokesperson.

3.6 Faculty Evaluations

3.6.0 Teaching Effectiveness

Each college defines what constitutes effective teaching through its faculty evaluation process. Components of teaching effectiveness may include but are not limited to:

- a. Performance in the classroom;
- b. Continuous updating, improvement, and innovation in teaching materials, methods, and assignments;
- c. Maintenance of regular office hours, at times convenient to students; and
- d. Advisement of students.

3.6.1 Faculty Evaluation Policy (SB)

Purpose -- The purpose of this document is to provide minimum standards for the evaluation of all full-time faculty. These procedures address evaluation as it relates to the development and the improvement of professional performance; in addition to the promotion, retention, and salary of those being evaluated.

3.6.1.0 Definitions

- a. Evaluation -- Evaluation is the process whereby the performance and competence of a person holding faculty rank are systematically examined and compared with established criteria.
- b. Position Description -- A position description is the written description of the scope and responsibilities of a position or group of positions held by faculty within the college.
- c. Criterion -- A criterion is the standard, rule, or test on which a judgment or decision can be based.
- d. College Plan -- A college plan is a detailed plan of evaluation prepared by each college in accordance with standards established by the State Board.

3.6.1.1 Application -- The procedures described herein shall apply to all full-time faculty.

3.6.1.2 College Plan

- a. Preparation of Plan -- Each college and the System Office shall prepare a detailed evaluation plan.

- b. Approval of Plan -- It is expected that all full-time faculty shall be involved in the development of the plan. The plan shall be approved by a majority of faculty and by the college president.
- c. Publication of Plan -- The college evaluation plan shall be included in the college's Faculty Handbook and a copy shall be transmitted to the office of the Chancellor.

3.6.1.3 Criteria -- The college evaluation plan shall include, but not be limited to, the following:

- a. Effectiveness in the performance of the tasks delineated in the appropriate position description;
- b. Effectiveness in establishing and maintaining positive professional relationships with colleagues, supervisors, students and the community;
- c. Effectiveness in maintaining a current competence in the particular discipline or field of specialization; and
- d. Adherence to policies, procedures, and regulations of the college and the VCCS.

Where additional criteria are considered, they shall be stated in the college plan.

3.6.1.4 Summary Ratings -- Performance evaluations shall include a summary rating of Excellent, Very Good, Good, Fair or Unsatisfactory as defined below:

- a. Excellent: consistently delivers outstanding performance, substantially exceeding performance standards.
- b. Very Good: clearly exceeds performance standards.
- c. Good: performs satisfactorily, meeting performance standards.
- d. Fair: marginally meets performance standards. Improvement required.
- e. Unsatisfactory: fails to meet performance standards.

3.6.1.5 Timetable -- The college plan shall contain a timetable that shall provide for completion of the evaluation process in time for the results to be used both in the development and improvement of professional performance; as well as the determination of promotion, retention, and salary.

- 3.6.1.5.0 Frequency -- The college plan shall require evaluation no less often than the following: (a) at least two times during the first year of employment and (b) at least one time during the second and each subsequent year of employment.
- 3.6.1.5.1 Notification -- The college plan shall provide that there shall be one or more conferences between the person being evaluated and the evaluator(s) at which time the results of the evaluation shall be discussed in detail. Moreover, the person being evaluated shall be provided a written summary of the evaluation.
- 3.6.1.5.2 Access to Records -- The college plan shall provide that the person being evaluated shall have the right to examine all materials utilized in the development of the evaluation. Faculty members shall be provided an opportunity to present a rebuttal, which shall become part of the record.
- 3.6.1.5.3 Appeal -- Administrative, professional, and teaching faculty may appeal their evaluation through the Faculty Grievance Procedure.
- 3.6.1.5.4 Review Process -- The college plan of evaluation shall be reviewed periodically. The review process shall provide for the involvement of all faculty. Recommendations for change shall be approved by a majority of the faculty and submitted to the president for final approval and implementation. If the recommended changes are not approved, the president must submit recommended modifications for further consideration and re-submission. In the meantime, the existing plan would remain in effect.

*VITA***WILLIAM H. HIGHTOWER, JR.**

Wytheville Community College
 1000 East Main Street
 Wytheville, VA 24382
 (276) 223-4794

EDUCATION

Ph.D. in Community College Leadership Old Dominion University	2010
M.S. in Zoology The Ohio State University	1981
B.A. in Biology The University of California, Los Angeles	1979

PROFESSIONAL HISTORY

Vice President of Instruction and Student Services Wytheville Community College, Wytheville, VA	2010 – Present
Director of Educational Programs Virginia Community College System Office, Richmond, VA	2006 – 2010
Dean of Instruction – Arts & Sciences, Business & Technology Southside Virginia Community College, Alberta, VA	1999 – 2006
Campus Coordinator – Virginia State University BRIDGES Project Southside Virginia Community College, Alberta, VA	1997 – 1999
Instructor – Anatomy & Physiology, Biology, Chemistry, Health Southside Virginia Community College, Alberta, VA	1992 – 1999
Adjunct Instructor – Anatomy, Anatomy & Physiology, Biology, Cross-sectional Anatomy	1991 – 1992
Engineering Designer – Civil Engineering Design, Drafting, Surveying Future Developments, Inc., San Diego, CA	1983 – 1991
Laboratory Preparator – General Biology Courses The Ohio State University, Columbus, OH	1981 – 1983
Graduate Teaching Associate – General Biology, Comparative Anatomy The Ohio State University, Columbus, OH	1979 – 1981

SELECTED PRESENTATIONS

“Cooperative, Multi-County Dual Enrollment Technology Programs” – VCCS New Horizons Conference, 2004

“Substituting Software for Scalpels – The A.D.A.M. Project” – VCCS State Board Annual Meeting, 1996

“Demonstration of Multimedia Technology for Classroom Instruction” – VCCS New Horizons Conference, 1994

ADVISORY BOARDS & TASK FORCE MEMBERSHIPS

Virginia Tech Engineering Education Advisory Board	2009 – 2010
VCCS Developmental Math Redesign Team	2009 – 2010
VCCS Sustainability Task Force	2009
Virginia Rural Health Resource Center Advisory Board	2008 – 2010
Geospatial Technician Education through Virginia’s Community Colleges (GTEVCC) Advisory Committee	2008 – 2010
VCCS Developmental Education Task Force	2008 – 2009
Governor’s Working Group on Early Childhood Initiatives	2007 – 2009
VCCS Faculty Evaluation Task Force	2007

LOCAL AND STATE ORGANIZATIONS

Chair, Council of Academic Deans	2004 – 2006
Leadership Mecklenburg Executive Committee	2004 – 2006
Assistant Director, Meherrin River District PTA	2002 – 2005
New Horizons Conference Committee	1998 – 2006
Chair, VCCS Sciences Peer Group Conference Committee	1996 – 1997

HONORS

Phi Kappa Phi Honor Society	2006 – Present
Who’s Who Among America’s Teachers	1996, 1998 & 2000

PUBLICATIONS

Hendrick, R. Z., **Hightower, W. H.**, & Gregory, D. E. (2006). State support of public higher education institutions and resulting limitations on continuation of the community college open door policy. *Community College Journal of Research and Practice*, 30, pp. 627-640.