

1-1-2004

Faculty Perceptions of the Effects of Student Evaluations of Teaching on Higher Education Instructional Practices and Instructor Morale

Annette Rashid Gall

Follow this and additional works at: <http://mds.marshall.edu/etd>

 Part of the [Curriculum and Instruction Commons](#), and the [Higher Education Administration Commons](#)

Recommended Citation

Gall, Annette Rashid, "Faculty Perceptions of the Effects of Student Evaluations of Teaching on Higher Education Instructional Practices and Instructor Morale" (2004). *Theses, Dissertations and Capstones*. Paper 598.

**FACULTY PERCEPTIONS OF THE EFFECTS OF STUDENT EVALUATIONS
OF TEACHING ON HIGHER EDUCATION INSTRUCTIONAL PRACTICES
AND INSTRUCTOR MORALE**

by

Annette Rashid Gall

**Dissertation submitted to
the Graduate College
of
Marshall University
in partial fulfillment of the requirements
for the degree of**

**Doctor of Education
in
Educational Leadership**

Approved by

**Dr. Teresa Eagle, Committee Chairperson
Dr. Michael Cunningham
Dr. Robert Rubenstein
Dr. John Moore**

Educational Leadership

ABSTRACT

FACULTY PERCEPTIONS OF THE EFFECTS OF STUDENT EVALUATIONS OF TEACHING ON HIGHER EDUCATION INSTRUCTIONAL PRACTICES AND INSTRUCTOR MORALE

by Annette Rashid Gall

The perceived consequences of student evaluations of teaching (SET) on higher education instructional practices and instructor morale were investigated. Participants were randomly selected from the 2002-2003 faculty of West Virginia's eight public colleges, n= 274. The researcher developed self-report survey, the *Gall Faculty Response to Evaluations of Teaching (FRET)*, was inspired by research literature. Research questions investigating the relationship between four independent variables and perception of changes in instructional practices were significant indicating that faculty consider information from student evaluations when changing teaching strategies. Research questions investigating the relationship between four independent variables and perception of changes in instructor morale were not significant indicating that the SET process has been institutionalized. Further research is needed to determine whether changes made to instructional practices increase or decrease learning and if the lack of effect on morale is positively or negatively related to learning.

ACKNOWLEDGMENTS

“Like everyone else, I own more debts than I can ever repay, or even tell”
(Holt, 1964. *How Children Fail*).

To those who have played a particularly important part in preparing me for my next 50 years-

-Dr. Teresa Eagle, for being the world’s best chair and mentor, now friend. You knew what I needed before I did. Thank you, thank you, thank you.

-Dr. Michael Cunningham, for generously sharing his plethora of expertise on topics ranging from designing questionnaires to designing buffets.

-Dr. Robert Rubenstein, for listening, guiding and inspiring me toward post-graduate education.

-Dr. Bobbi Nicholson, Marshall’s first Fulbright Scholar, my dear friend, for refreshing the liberal dreams of my youth and for polishing the analytic skills by which those dreams achieve reality.

-Dr. Rebecca Goodwin, generous friend and mentor, for organizing what I couldn’t see. Teaching with you is a joy.

-Dr. Dennis Prisk who warned me about sexism in higher education and inspired me to conquer it.

-Double Dr. Judy Porter, the sister I found in the program, who loves me anyway I am.

All of you- Teresa, Mike, Bob, Bobbi, Becky, Dennis, and Judy- all of you are consummate educators and now my colleagues. Thank you.

To extended family and friends, who cheered me along-

-My mom and my sister Lenore, who fostered in me their dreams of education and modeled the value and practice of life-long learning.

-John Moore, for unwavering support for 39 years.

-My brother-in-law Jim Hendricks, consummate journalist and editor, for teaching me so much.

-Megan and Chris, who grew from children to adults while I reinvented myself.

-My husband Chris, my hero, who gave me the world and makes all things possible.

Thank you all. You are a part of me, whether you like it or not.

TABLE OF CONTENTS

ABSTRACT.....	iii
ACKNOWLEDGMENTS	iv
TABLE OF CONTENTS.....	v
LIST OF TABLES.....	viii
CHAPTER I	1
INTRODUCTION.....	1
BACKGROUND.....	2
INTENDED PURPOSES OF STUDENT EVALUATIONS OF TEACHING.....	2
Formative Student Evaluations of Teaching for Instructional Improvement.....	4
Summative Student Evaluations of Teaching for Faculty Promotions.....	4
CONSEQUENCES OF STUDENT EVALUATIONS OF TEACHING.....	6
The Effects of Student Evaluations of Teaching on Instructional Practices.....	6
The Effects of Student Evaluations of Teaching on Instructor Morale.....	7
FACULTY OPINIONS OF STUDENT EVALUATIONS OF TEACHING.....	8
Faculty Opinions of the Validity of Students as Evaluators.....	9
Faculty Opinion about the Appropriate Uses of Student Evaluations of Teaching.....	10
Faculty Opinion of Their Ability to Improve Student Evaluations of Teaching Ratings.....	10
FACULTY USE OF STUDENT EVALUATIONS OF TEACHING	11
EXPECTANCY THEORY.....	12
STATEMENT OF PROBLEM.....	14
RESEARCH QUESTIONS.....	14
OPERATIONAL DEFINITIONS.....	16
SIGNIFICANCE TO HIGHER EDUCATION ADMINISTRATORS.....	17

LIMITATIONS OF STUDY.....	19
CHAPTER II.....	21
REVIEW OF RELATED LITERATURE.....	21
INTRODUCTION.....	21
BACKGROUND.....	23
PURPOSES OF STUDENT EVALUATIONS OF TEACHING.....	23
The Development of Student Evaluations of Teaching in American Education.....	24
The Purposes of Student Evaluations of Teaching.....	25
Formative Student Evaluations of Teaching.....	27
Summative Student Evaluations of Teaching.....	29
CONSEQUENCES OF STUDENT EVALUATIONS OF TEACHING.....	31
The Effects of Student ?Evaluations of Teaching on Instructional Practices.....	31
The Effects of Student Evaluations of Teaching on Instructor Morale.....	33
FACULTY OPINIONS OF STUDENT EVLAUATIONS OF TEACHING.....	34
Faculty Opinion of the Validity of Students as Evaluators.....	35
Faculty Opinion about the Appropriate Uses of Student Evaluations of Teaching	36
Faculty Opinion of Their Ability to Improve Student Evaluations of Teaching.....	37
FACULTY USE OF STUDENT EVALUATIONS OF TEACHING	38
EXPECTANCY THEORY.....	40
CHAPTER III.....	43
METHODS.....	43
POPULATION AND SAMPLE.....	44
INSTRUMENTATION.....	45
METHODS.....	46
DATA ANALYSIS.....	47
SUMMARY.....	48

CHAPTER IV.....	49
STATISTICAL ANALYSIS OF THE DATA.....	50
DESCRIPTIVE DATA.....	55
ANCILLARY FINDINGS.....	60
SUMMARY.....	61
CHAPTER V.....	62
METHODS.....	62
Instrumentation and Data Analysis.....	63
DESCRIPTIVE DATA.....	63
FINDINGS AND CONCLUSIONS.....	64
Ancillary Findings.....	71
IMPLICATIONS TO HIGHER EDUCATION ADMINISTRATORS.....	72
GENERALIZABILITY.....	76
RECOMMENDATIONS FOR FURTHER STUDY.....	77
BIBLIOGRAPHY.....	80
APPENDICES.....	87
A. GALL FACULTY RESPONSE TO EVALUATIONS OF TEACHING (FRET)	87
B. FIRST COVER LETTER.....	90
C. REMINDER POSTCARD.....	92
D. FOLLOW-UP COVER LETTER.....	94

LIST OF TABLES

Table 1. Faculty Perceptions of the Relationship of Student Evaluations of Teaching on Instructional Changes and Instructor Morale.....	51
Table 2. Frequency Distribution by Faculty Status.....	55
Table 3. Frequency Distribution for College/School.....	56
Table 4. Frequency Distribution of Ethnicity.....	57
Table 5. Frequency Distribution for Years in Higher Education.....	58
Table 6. Frequency Distribution for Age.....	59
Table 7. Frequency Distribution for Years at Present Institution.....	60

CHAPTER I

Introduction

Though the actual process may vary from one postsecondary institution to another, student evaluations of teaching (SET) are the most frequently used instruments for evaluating higher education faculty (Seldin, 1984; Spencer & Flyr, 1992). Student evaluations are probably "...the most thoroughly studied of all forms of personnel evaluations and one of the best in terms of being supported by empirical research" (Marsh, 1987, p. 369). Abundant literature exists on the composition of the SET and reasons for using it (Read, Rama, & Raghunandan, 2001; Wachtel, 1998; Wallace & Wallace, 1998). A wealth of research is available on the method's validity and reliability (Aleamoni, 1987; Cashin, 1988; Marsh, 1984). The potential biases of teacher characteristics, course characteristics, student characteristics and environmental characteristics have been established (Calderon & Green, 1997; Martin, 1998).

Missing in the literature is research on the effect of student ratings on instructional practices and morale (Ryan, Anderson, & Birchler, 1980; Simpson & Siguaw, 2000; Spencer & Flyr, 1992; Stratton & Myers, 1994; Wachtel, 1998). Are the effects beneficial and energizing or intimidating and threatening? Perhaps the explanation for this omission in research literature lies in the difficulty in obtaining baseline data about teaching evaluations before the SET was institutionalized (Stratton & Myers, 1994). This study will investigate faculty perceptions of the effects of student evaluations of teaching on instructional practices and instructor morale using Expectancy Theory to predict the faculty member's response.

Background

Numerous sources have criticized American schooling for its role in educating students (Carnegie Forum on Education and the Economy, 1986). The American Association for Higher Education and other organizations recently emphasized that universities and colleges need to reassess their commitment to instruction (Read, et al, 2001) and place greater importance on teaching (Perry & Smart, 1997). “Scholarship is not an esoteric appendage: it is at the heart of what the profession is all about...to weaken faculty commitment for scholarship is to undermine the undergraduate experience” (Boyer, 1990, p. 1).

Intended Purposes of Student Evaluations of Teaching

Seldin (1980) stated that the main purpose of the SET was to improve teaching by moving faculty toward excellence. Seldin (1984) further explained that “The cornerstone of every faculty evaluation program is its purpose: the purpose influences the kinds of questions asked, sources of data, depth of data analysis, and dissemination of findings” (p. 127). In 1976, the Southern Regional Educational Board described faculty evaluations as having two purposes, formative and summative (Centra, 1993). Modern SET practices were originally designed as a benign tool to be used formatively in support of faculty development and self-improvement (Simpson & Siguaw, 2000; Stapleton & Murkison, 2001; Trout, 1997; Zelby, 1974). Now they have a summative purpose as well: viewed as a convenient and ostensibly objective measure of teaching efficacy, they are employed in administrative decisions on salary and promotion (Stapelton & Murkison, 2001; Trout, 1997).

Student evaluations of teaching provide numerical evaluations of instruction, allegedly relieving administrators from judging teaching performance and ability (Wilson, 1998; Zelby, 1974). However, using student evaluations summatively has changed them from a helpful guide for professional growth into a mechanism that could harm the evaluation process (Zelby, 1974). According to Centra (1993), the only appropriate use for formative evaluations is as an aid to improving instruction; summative use of the same evaluations changes the effects on the instructor and the role of the evaluator, perhaps to the detriment of the formative purpose.

In higher education, the principal purposes of student evaluations are to improve performance and to provide rationale for administrative decisions on tenure, promotion and retention (Seldin, 1984). Ideally, faculty evaluations for purposes of improving teaching would be conducted separately from evaluations for personnel decisions (Seldin, 1984). For improving teaching, ratings would provide an accurate diagnosis of problems leading to potential solutions. If the intended use were for personnel decisions, ratings would be valid measures of teaching effectiveness. For student guidance, ratings would provide valid information allowing students to choose the more valuable learning experience (McKeachie, 1979). Given time and fiscal restraints, however, it is doubtful that higher education can conduct separate ratings (Seldin, 1984). Therefore, it becomes important to also confirm the effects of using the SET for both formative and summative reasons on instructional practices and morale.

Formative Student Evaluations of Teaching for Instructional Improvement

Evaluations of teaching have existed in one form or another since the time of Socrates (Marsh, 1987). Today, the SET is an essential process on which higher education administrators rely to assess the intangible concept of instructional quality in a quantitative, concrete and reportable manner (Simpson & Siquaw, 2000). The use of SET is prevalent in higher education institutions because they are widely available, are seemingly quantitative in nature and provide a comparable basis of data collection among institutions (Wallace & Wallace, 1998). Greenwald and Gilmore (1997) predicted that the SET will continue to be used extensively because there is no available alternative procedure for evaluating instruction. Additionally, the recent public scrutiny and criticism of professors has made teaching quality a priority for professional school administrators (Frost & Fukami, 1997).

Foote (1998) asserted that maintaining a high standard of excellence requires performance evaluations of faculty and Wilson (1998) added that the SET is sometimes the only measure of teaching ability. This is significant as postsecondary institutions are encouraged to revise policies on tenure, promotion and merit to give teaching greater emphasis (Perry & Smart, 1997). Revising instructional evaluation procedures would also stand as proof that teaching is important to higher education institutions (Simpson & Siquaw, 2000; Zelby, 1974).

Summative Student Evaluations of Teaching for Faculty Promotion

In 1984, Seldin predicted the then current and unprecedented financial strain on higher education would continue to worsen. Adding to fiscal constraints was the expectation of diminished numbers of college applicants causing intense competition for students

between higher education institutions, in addition to further rivalry from corporate universities (Seldin, 1984). “In an era of growing accountability and outcomes evaluations, achieving a better understanding of the evaluation of teaching effectiveness may be a necessary step toward including scholarship of teaching in decisions on faculty tenure and promotion” (Hobson & Talbot, 2001, p.34).

Higher education administrators have assigned importance to the SET in the past 20 years because of the increased competition for faculty positions and the attempt to find an impartial basis on which to appraise faculty for reappointment, tenure and promotion (Williams & Ceci, 1997). While recognizing the institutionalization of the ratings, however, Stapleton and Murkison (2001) questioned whether the use of student evaluations has resulted in grade inflation and decreased academic standards. Trout (2000) concurred that the SET is deeply entrenched in postsecondary institutions and stressed that he and others involved in the evaluation process consider numerical forms used to reward and punish as a detriment to instructional practices. Changing the purpose of the SET to fulfill the need for faculty appraisals could have made the process harmful to teaching (Zelby, 1974).

Out of the financial strain that Seldin (1984) predicted came considerable changes for higher education. The student-oriented or student-as-a-customer approach to education was initiated (Simpson & Siguaw, 2000), and contributed to accusations that universities now pander to students by diluting instruction and inflating grades to acquire higher customer satisfaction (Swenson, 1999; Wilson, 1998). During this same period, state governments and the public began holding higher education answerable to measurable outcomes by allocating resources using performance indicators as a guide for appropriations (Simpson & Siguaw, 2000).

Using the SET for both formative and summative purposes placed pressure on faculty members. They have found themselves struggling to maintain high standards of instruction while, at the same time, working to elicit favorable ratings of their performance from students who may be more concerned with grades than with learning.

Consequences of Student Evaluations of Teaching

Research has not yet included changes in instructional practices and instructor morale as a result of the use of SET (Ryan, et al, 1980; Spencer & Flyr, 1992; Stratton & Myers, 1994; Wachtel, 1998). Early research provided rationale for or against using the SET and focused on the properties, characteristics and content of evaluation instruments or the effects of various conditions for obtaining information (Kulik & McKeachie, 1976).

The Effects of Student Evaluations of Teaching on Instructional Practices

Evidence that student ratings result in improved teaching is mixed (Seldin, 1980). Although some researchers have contended that using SET information did not bring improvement (Centra, 1972; Miller, 1971), others found improvement under certain circumstances (Gage, 1972; McKeachie, 1972; Wilson, 1998). In a review of the literature, Seldin (1980) concluded that hard evidence about improved classroom teaching is lacking. Millman (1981) agreed, stating, “Although there has been a great deal of anecdotal evidence from instructors and researchers to suggest that student evaluations do have a positive effect, very few studies are available that deal with that effect on college-level instruction” (p. 140).

In the modern seminal research on the influence of the SET on faculty, Ryan et al. (1980) reported that mandatory evaluation programs have decreased the rigor of classroom

teaching. Their study demonstrated a definite increase in instructional practices judged counter-productive to education's mission. Other researchers concurred, finding that good evaluations may be associated with teaching of lesser educational value (Trout, 1997; Wallace & Wallace, 1998; Zelby, 1974)

Since using evaluations administratively makes the process economically important to faculty, Simpson and Siquaw (2000) suspected some faculty may employ practices that affect the SET process rather than improve teaching (Krautman and Sanders, 1999). Trout (1997) stated:

It is hard to imagine a practice more harmful to higher education than one that encourages instructors to satisfy the demands and pleas of students who resent the appropriate rigors of college instruction. These forms are not just invalid and unreliable; they are pernicious (p. 30).

The Effects of Student Evaluations of Teaching on Instructor Morale

The research literature on the effects of morale on teaching quality is sparse. Ryan et al. (1980) reported a large portion of responding faculty experienced significant reduction in staff morale and reduced job satisfaction as a result of SET information, a finding supported by industrial research. Employees receiving less than outstanding ratings may experience a significant drop in organizational commitment (Pearce & Porter, 1986), as well as becoming apprehensive about the organization (Gabris & Mitchell, 1988). Other researchers reported a direct relationship between performance ratings, overall job satisfaction and future performance (Blau, 1999; Kohli & Jaworski, 1994).

Faculty vitality is at the heart of the institution's fundamental nature (Des Jarlais, 1996). According to the Faculty Professional Self-Esteem paradigm, teachers are more effective when morale is high and they feel respected and valued (Halford, 1994). Further research into issues that influence faculty morale seems meaningful (Ryan et al., 1980) because raising faculty morale has been identified as one of the crucial jobs of higher education administration (Des Jarlais, 1996). In addition, more research is needed to gauge faculty member's views of SET application and its impact on their teaching content and style.

Faculty Opinions of Student Evaluations of Teaching

Despite differences in opinions and research findings, the SET is primarily and nearly universally, used as a measure of teaching effectiveness appraising the quality of faculty on selected instructional dimensions (Hobson & Talbot, 2001; Marsh, 1987; Simpson & Siquaw, 2000). Student evaluations of teaching are the current standard and the merits or weaknesses are not debated in his study. The prolific research on the SET, however, has not addressed faculty opinions about student evaluations which might affect their perception of the information communicated through evaluations (Ryan et al, 1980; Spencer & Flyr, 1992; Stratton & Myers, 1994; Wachtel, 1998). This section explores faculty opinions of (a) students as evaluators, (b) appropriate uses of the SET, and (c) faculty ability to improve SET ratings.

Faculty Opinions of the Validity of Students as Evaluators

Remmers (1928) maintained the following: “Should it be true that the average student tends to have his attitudes toward instruction influenced by the marks he receives, this would constitute an important psychological fact to be kept in mind by the instructor and possibly by the student” (p. 759). In 1975, McKenzie provided support that the SET reflect the degree to which the student and instructor agree on factors such as grades, content and classroom entertainment. Wallace and Wallace (1998) argued that the SET measures the students’ happiness at the end of the course before grades are received.

Some have argued that students are the best evaluators for teaching. “It is manifestly true that the only direct, daily observers of a professor’s classroom teaching performance are the students in the classroom” (Seldin, 1980, p. 36). Marsh (1998) contended that as higher education is organized, student raters have spent more time observing instruction than any other potential raters. Feldman (1997) agreed and stated that students as raters of instruction are an obvious and pragmatic choice.

Other researchers reported that teaching ratings are not accurate measures of teaching performance (Gomez-Mejia & Balkin, 1992; Simpson & Siguaw, 2000). As the scope of the SET widened, some maintained that subject matter was included that exceeds what one would expect a student to be qualified to rate (Wallace & Wallace, 1998).

Human resource researchers, investigating relationships similar to the student/instructor dyad, maintained that only some aspects of a manager’s work are appropriate for evaluation by subordinates (Bernardin, Dahmus & Redmon, 1993). McEvoy (1990) confirmed that subordinates may appropriately evaluate a manager’s skills in

communication, interpersonal relations management, and leadership. Inclusion of other aspects may undermine confidence in the ratings (Bernardin, Dahmus & Redmon, 1993).

Faculty Opinion about the Appropriate Uses of Student Evaluations of Teaching

Many authors reported indecisions among faculty about the appropriateness of the use of the SET for personnel and tenure decisions (Rutland, 1990; Ryan et al., 1980; Zoller, 1992). Rifken (1995) maintained that the inability to agree on the intended purposes of evaluations is one of the main obstacles to developing more effective assessments. In 1974, Zelby warned that unconsidered use of SET, particularly for determining promotions and tenure, would inhibit experimentation and development of education.

Some question whether one instrument can serve both formative and summative purposes, i.e., correctly diagnosing teaching problems and proposing solutions while simultaneously being used for deciding faculty merit, promotion and tenure (Rifkin, 1995; Simpson & Siguaw, 2000). Simpson and Siquaw (2000) concluded that use of the SET for summative purposes makes the instrument vitally important to faculty, perhaps encouraging lowering of educational standards.

Faculty Opinion of Their Ability to Improve Student Evaluations of Teaching Ratings

Simpson and Siguaw (2000) found some instructors instituted practices specifically designed to increase SET ratings such as grade inflation, “dumbing down” of coursework and lowered workload (Ryan et al, 1980; Zelby, 1974). Such practices are, of course, counter to the educational purpose of learning (Baxter, 1991; Simpson & Siguaw, 2000; Zelby, 1974). Confounding the issue, typical evaluation approaches cannot differentiate between instructors

who raise students' grades by increasing learning efficiency and instructors who raise grades via lower class requirements (Stapleton & Murkison, 2001; Stratton & Myers, 1994).

Ryan et al. (1980) reported that a majority of faculty respondents believed SET ratings were less influenced by academically relevant achievement and more influenced by irrelevant personal qualities of the faculty. A substantial proportion of the respondents believed they could not improve their SET ratings and were not likely to try (Ryan et al., 1980). Franklin and Theall (1989), however, concluded that the anecdotal evidence of faculty opposition towards the SET is extensive. Simpson and Siguaw (2000) concurred that instructors perceive problems with student evaluation ratings and suggested that the next step in research should examine the pervasiveness of disenchantment with the SET and the mode of response instructors choose to influence SET ratings.

Faculty Use of Student Evaluations of Teaching Results

Despite the evidence of faculty opposition to the SET process and after several decades of research, the majority of researchers accept that the SET is valid, reliable, and worthwhile (Centra, 1993; McKeachie, 1990; Seldin, 1993). Franklin and Theall (1989) proposed that faculty familiar with the SET literature would make the best use of ratings. They anticipated that positive attitudes towards the SET process should be associated with better practice, and that those who are ignorant of the facts of evaluations are at greater risk of carrying out bad practices. Franklin and Theall (1989) emphasized that faculty may not know what is needed to make the best use of SET results, stating "...there is insufficient evidence to conclude that the results are understood and applied by users with at least the validity and reliability of the instruments that obtained them" (p. 1). Faculty appeared not to

be aware of current literature that might help them make informed decisions about appropriate uses of the SET. Lack of awareness may lead instructors to expect that the SET will have harmful impact. Under Expectancy Theory, such negative expectations prevent faculty from viewing SET favorably and gaining the best possible benefit from it.

Expectancy Theory

All renditions of Expectancy Theory have roots in Vroom's Expectancy Theory of Motivation which attempted to provide an explanation of how people choose between behaviors (Vroom, 1964). Vroom hypothesized that the motivation and effort to perform is a function of the probability that the requisite behavior will result in the desired outcome (Vroom, 1964). Behavior is the result of selecting from alternatives that will maximize pleasure, minimize pain and produce the best rewards. Expectancy Theory addresses an individual's motivation to behave in a certain way in a particular situation, rather than an individual's overall motivation (Ratzburg, 2002).

When choosing between behaviors, individuals select the option that has the highest motivational force (Ratzburg, 2002). Motivational force is the product of the three perceptions of expectancy, instrumentality and valence. Expectancy is the strength of an individual's belief that the job can be accomplished, and that the individual's effort will result in performance goals (Vroom, 1964). The individual's belief is based on the level of self-confidence, past experiences and perceived difficulty of the performance goal (Vroom, 1964). Instrumentality is the perceived probability that a high level of performance will result in desired outcomes, that performance is instrumental in acquiring rewards (Vroom, 1964). To insure high levels of performance, management must tie desired outcomes to

performance and ensure that individuals understand the connection. Both expectancy and instrumentality represent the individual's perception of the likelihood that the required effort will result in the desired outcomes.

Valence is subjective and refers to the value of the expected rewards to the individual or the level of satisfaction the individual expects from the outcome. People select the level of performance that will best meet their needs, goals and values (Top Education, 2001).

In this study, expectancy relates to whether instructors perceive that their efforts will lead to improved student learning and/or promotion, tenure and raises. Instrumentality is related to whether instructors believe that improved student achievement will earn them a boost in morale, tenure, promotion or a raise. Valence is indicative of whether instructors individually value or find desirable positive consequences, like satisfaction of seeing students succeed or personal career rewards. Valence can also indicate whether the individual instructor wants to avoid negative repercussions associated with not meeting student achievement goals (Odden, 2000).

Motivation will be low if individuals do not believe they can be successful at the required task, if they believe that a successful task will not lead to positive outcomes, or if they believe outcomes will be negative. If instructors perceive that earning higher SET ratings is too difficult or that they cannot influence SET ratings, student comments may be ignored and instructors' morale may be low. Additionally, if instructors do not believe that the goals of promotion and tenure are possible, morale is also low.

The motivation to improve teaching to increase SET ratings depends on the importance of the outcomes and the degree to which higher ratings are believed to result in desired outcomes such as promotion, tenure and raises. According to Expectancy Theory,

SET ratings should become more important to faculty when the evaluations are used to help decide promotions or are used altruistically to increase student learning. If these rewards are important to instructors, Expectancy Theory asserts that faculty will exert substantial effort to increase their SET ratings.

Statement of the Problem

As postsecondary institutions reconsider policies on tenure, promotion and merit to give teaching greater emphasis (Perry & Smart, 1997), higher education administrators have come to rely on the SET to assess the intangible concept of instructional quality in a quantitative and reportable manner (Simpson & Siguaw, 2000). In light of current budget cuts and demands for accountability, researching the perceived effects of evaluations on instructional practices and instructor morale is warranted. The literature indicates that researchers have not sufficiently explored the effects of several faculty beliefs. Specifically, this investigation will collect information about faculty opinion of the validity of students as evaluators, faculty opinion about the appropriate uses of SET, faculty opinion of their ability to improve SET ratings and faculty use of SET feedback. This study will add to the body of knowledge about faculty perceptions of the effects of student evaluations of teaching on instructional practices and instructor morale.

Research Questions

The purpose of this study was to determine the perceived effects of student evaluations of teaching on instructional practices and instructor morale. The population was members of the faculties of the eight public, four-year colleges in West Virginia (N =

approximately 935) during the 2002-2003 academic year. An appropriately sized ($n =$ approximately 274) random sample was chosen from this population to participate in the study (Johnson & Christensen, 2000). The *Gall Faculty Response to Evaluations of Teaching (FRET) Survey*, designed by the researcher, was used to gather data related to the following questions:

1. Is there a statistically significant relationship between faculty opinion of the validity of students as evaluators of teaching and faculty perceptions of changes in instructional practices?
2. Is there a statistically significant relationship between faculty opinion of the validity of students as evaluators of teaching and faculty perceptions of changes in morale?
3. Is there a statistically significant relationship between faculty opinion of the appropriate uses of student evaluations of teaching and faculty perceptions of changes in instructional practices?
4. Is there a statistically significant relationship between faculty opinion of appropriate uses of student evaluations of teaching and faculty perceptions of changes in morale?
5. Is there a statistically significant relationship between faculty perceptions of their ability to influence ratings from student evaluations of teaching and faculty perceptions of changes in instructional practices?
6. Is there a statistically significant relationship between faculty perceptions of their ability to increase ratings from student evaluations of teaching and faculty perceptions of changes in morale?

7. Is there a statistically significant relationship between faculty interest in student evaluations of teaching and faculty perceptions of changes in instructional practices?
8. Is there a statistically significant relationship between faculty interest in student evaluations of teaching and faculty perceptions of changes in morale?

Operational Definitions

For the purpose of this study, the following operational definitions are used:

1. Perceived instructional changes- a faculty member's response to the instructional changes components of the researcher designed *Gall Faculty Response to Evaluations of Teaching (FRET) Survey*.
2. Instructor morale- a faculty member's response on the morale components of the researcher designed *Gall Faculty Response to Evaluations of Teaching (FRET) Survey*.
3. Opinion of validity of students as evaluators- a faculty member's response to the students as evaluators components of the researcher designed *Gall Faculty Response to Evaluations of Teaching (FRET) Survey*.
4. Opinion of appropriate uses of the SET- a faculty member's response to the appropriate uses components of the researcher designed *Gall Faculty Response to Evaluations of Teaching (FRET) Survey*.
5. Opinion of faculty ability to increase SET - a faculty member's response to the ability to increase SET components of the researcher designed *Gall Faculty Response to Evaluations of Teaching (FRET) Survey*.

6. Interest in the SET- a faculty member's response on the interest components of the researcher designed *Gall Faculty Response to Evaluations of Teaching (FRET) Survey*.

Significance to Higher Education Administrators

Public confidence in colleges and universities declined between 1965 and 1985 (Johnsrud, 1996). Since then, economic and societal conditions have worsened and higher education has not found solutions to these issues to satisfy the public (Johnsrud, 1996). The American Association for Higher Education and other organizations have called for colleges and universities to reassess their commitment to teaching and place greater importance on it (Read et al., 2001). Boyer (1990) believed that while professors are often caught between competing duties, teaching is, and must remain, the primary task of higher education.

Higher education presidents and their administrative staffs have many new challenges to meet in managing the damaged credibility that has influenced morale and led to legislative funding cuts (Johnsrud, 1996). Reviewing teaching evaluation policies would demonstrate that higher education administration is committed to improving teaching (Simpson & Siquaw, 2000).

As supporting faculty morale is one of the crucial functions of higher education administration that affects an institution's substance, tone and momentum (Des Jarlais, 1996), presidents may apply the information from this study when organizing and developing the strategic plan for improving instruction and raising faculty morale. Johnsrud (1996) viewed morale as part of the evaluation process, and Halford (1994) contended that teachers are more effective when they feel valued and respected. Certainly the highest goals

for the institution deserve the president's full understanding of factors affecting instruction when coordinating the institution's activities.

The chief administrative officer may employ the results of the study to budget and coordinate the levels of financial support allocated for faculty development and morale building programs, as well as support for the SET program itself. Funding programs that maintain or enhance educational quality will serve to strengthen the institution. Financial officers can better preserve institutional assets if the highly competitive market perceives the institution as a center of academic excellence thus attracting capable students.

The chief academic officer can use the information from this study to direct the choice of relevant development and orientation programs for faculty. Faculty development strengthens faculty morale, improves teaching effectiveness and increases job satisfaction, all of which contribute to the institution's energy (Kang & Miller; 2000).

Renewed emphasis on teaching outcomes and accountability makes teacher evaluation programs increasingly significant in the reorganization of staff and budgeting policies regarding promotion, tenure and merit (Feldman, 1997; Perry & Smart, 1997; Read et al., 2001; Waters et al., 1988). Hobson and Talbot (2001) emphasized that a better understanding of evaluations of teaching effectiveness may be required if teaching is to become more important in decisions concerning faculty tenure and promotion.

The academic chairperson may use the information from this study to better appreciate the faculty response to the SET and thereby more effectively support staff in personal and professional development. Understanding the faculty response to the SET process may allow the chairperson to improve communication and coordinate performance expectations and priorities for faculty members.

Understanding the SET process would also benefit the individual faculty member and support a positive faculty response to SET ratings. Faculty members can learn to analyze student comments to plan and organize class content to better communicate with students.

Student evaluations of teaching are sometimes the single measure of teaching aptitude (Wilson, 1998) on which higher education administrators rely to assess the complexities of instructional quality (Simpson & Siquaw, 2000). The SET process has become important in the past 20 years because of the increasing search for an impartial faculty appraisal process (Williams & Ceci, 1997). As higher education institutions experience budget cuts, demands for accountability and renewed interest in teaching, researching the perceived effects of evaluations on instructional practices and instructor morale is warranted.

Limitations of Study

1. Data in this study are from faculty members of four-year public colleges in West Virginia and the results may not generalize to faculty members in other types of institutions or in other states or countries (Kerlinger & Lee, 2000).
2. This study does not account for other faculty attributes, such as personality or teaching style, which might affect their perceptions of the SET.
3. The measurements of the faculty member's instructional changes and morale due to the SET are the perceptions of the faculty member.
4. The measurements of the faculty member's interest in the SET and opinions of student evaluators, appropriate uses and ability to increase ratings are the perceptions of the faculty member.

5. The validity of the *Gall Faculty Response to Evaluations of Teaching (FRET)* Survey instrument will present limitations on the results of the study (Kerlinger & Lee, 2000).
6. The study employs a self-reported questionnaire survey and is limited by the accuracy of the faculty member's responses (Kerlinger & Lee, 2000).

CHAPTER II

Review of Related Literature

This chapter presents a review of literature associated with faculty perceptions of the effects of student evaluations of teaching (SET) on instructional practices and instructor morale. The dependent variables for this study are the effects of SET ratings on instructional practices and instructor morale. The independent variables are faculty opinion of the validity of students as evaluators, faculty opinion of the appropriate uses of the SET, faculty opinion of their ability to improve SET ratings and perceived faculty interest in SET information. Demographic information was also solicited. The population for this study was members of the faculties of eight public, four-year colleges in West Virginia. Data was solicited from a random sample of the population.

The research literature contains an abundance of evidence concerning validity, reliability and many confounding factors. However, there is a dearth of empirical evidence about the effects of SET ratings on instructional practices and instructor morale. This study added to the body of knowledge concerning what types of changes in teaching practices faculty make after considering SET information and how SET information affects their morale.

Introduction

Student evaluations of teaching are the most commonly used instrument for higher education faculty evaluation (Seldin, 1984; Spencer & Flyr, 1992), with the vast majority of colleges and universities in the United States requiring some form of evaluation by students (Wachtel, 1994). Carnegie Foundation for the Advancement of

Teaching reports that 98% of universities use SET information (Manger, 1997) as sometimes the only measure of teaching ability (Wilson, 1998). An enormous amount of research literature exists on SET (Marsh & Dunkin, 1987; Read, Rama, & Raghunandan, 2001; Wachtel, 1998; Wallace & Wallace, 1998) with probably more studies than all other means of evaluating college faculty combined (Calderon & Green, 1997; Cashin, 1988; Wallace & Wallace, 1998). Wilson (1998) reported that 2000 studies were conducted on student evaluations since the 1920s. Marsh (1987) wrote that SET are probably "...the most thoroughly studied of all forms of personnel evaluation, and one of the best in terms of being supported by empirical research" (p. 369).

Heavy reliance on SET for evaluating teaching effectiveness raises questions about validity, reliability and the effects of extraneous variables on student responses (Simpson & Siquaw, 2000). Reviews of SET literature consistently show that they are multidimensional, reliable, relatively valid, useful for improving teaching and relatively unaffected by confounding factors (Aleamoni, 1987; Cashin, 1988; Marsh, 1984; Marsh & Roche, 2000). The list of well researched confounding factors includes teacher characteristics, course characteristics, student characteristics and environmental characteristics (Calderon et al., 1996; Martin, 1998). However, research on the potential effects of SET ratings on instructional practices and instructor morale is lacking in the literature (Ryan, Anderson, & Birchler, 1980; Simpson & Siquaw, 2000; Spencer & Flyr, 1992; Stratton & Myers, 1994; Wachtel, 1998).

In 1976, Kulik and McKeachie questioned whether the effects of SET on faculty and teaching were refreshing and healthy or demoralizing and dangerous. In 2001, Stapleton and Murkison expressed concern that the use of SET resulted in lowered

academic standards. Research such as this study into the perceptions of faculty members may serve as an impetus to changing how teaching is measured and how SET are used in performance evaluations to encourage positive faculty response (Simpson & Siquaw, 2000).

Background

Between 1965 and 1985, public confidence in postsecondary institutions diminished. The public questioned the ability of campus leaders to control student unrest in the late 1960s. As societal and economic conditions worsened, the public felt higher education was not providing adequate solutions (Budig, 1986). The Carnegie Forum on Education and the Economy (1986) documented that American education was being censured for its role in educating students (Lanning & Perkins, 1995). The American Association for Higher Education and other organizations suggested that postsecondary institutions reassess their commitment to teaching (Read, Rama, & Raghunandan, 2001) and assign greater importance to teaching (Perry & Smart, 1997). The loss of credibility had a direct effect on the morale of those in higher education and allowed state legislators to cut funding to colleges and universities (Budig, 1986).

Purposes of Student Evaluations of Teaching

Socrates may have been among the first teachers held accountable by evaluations; he was executed in 399 BC for allegedly corrupting the youth of Athens with his teachings (Marsh, 1987). Teacher evaluations can be traced to AD 350 in Antioch when fathers dissatisfied with their sons' education filed a formal complaint against the teacher

(Marsh, 1987). In medieval Europe, student committees monitored the information instructors covered, reported irregularities for which instructors were fined and determined instructors' salaries by considering how many students attended their classes (Centra, 1993). By the early 1800s in America, lay committees inspected Boston schools to determine if educational objectives were met (Spencer & Flyr, 1992).

The Development of Student Evaluations of Teaching in American Education

In the 20th Century, modern student evaluation programs began in the United States at several major universities (Wachtel, 1998). In 1924, Harvard students collected and printed student ratings of instructors in the *Confidential Guide to Courses* that inventoried student opinions of courses and professors (McKeachie, 1979; Seldin, 1980). Other schools introducing these first student evaluations include University of Washington, Purdue University and University of Texas (Radmacher & Martin, 2001).

The modern era of evaluations may be divided into four periods: (a) 1927-1960 (b) 1960s, (c) 1970s, and (d) 1980s to the present (Centra, 1993). During the first period, H. H. Remmers and his colleagues established the first systematic examination into student evaluations of teaching efficacy at Purdue University (Carroll, 2002; Centra, 1993; Marsh, 1987). During the 1960s, teaching evaluations were almost entirely voluntary (Wachtel, 1998). However, the explosive development of the SET is credited to the campus unrest of that decade (Seldin, 1980; 1984; Wilson, 1998), characterized by student protests concerning the Vietnam War, governmental policies and campus policies (Centra, 1993). Students demanded a voice in their education, either by organizing their own evaluation system or by pressuring institutions to develop evaluation instruments

(Centra, 1993). Under the stress, the nature of the professoriate changed from an intellectual calling to a job requiring evaluations to meet the demands for public accountability (Seldin, 1984).

Interest in validity, bias and utility of student evaluations encouraged intense research in the 1970s (Centra, 1993). By the close of the decade, most institutions used student evaluations as the most important source for assessing instructional effectiveness, fulfilling higher education's administrative need for objective data for making bureaucratic decisions (Centra, 1993; Seldin, 1984). Inquiry from 1980 to 1993 centered on interpretation and refinement of the SET instrument (Centra, 1993).

The Purposes of Student Evaluations of Teaching

Today, higher education administrators depend on SET information to appraise the elusive idea of instructional quality in a quantitative, concrete and reportable method (Simpson & Siquaw, 2000). Higher education institutions use the SET process because it is generally accessible, apparently quantitative and offer an equivalent basis of information collection among institutions (Wallace & Wallace, 1998). With no efficient alternative for evaluating teaching, the SET is likely to remain the most extensively used method of teaching evaluation (Greenwald & Gilmore, 1997).

Faculty performance evaluation is essential to sustain a high standard of excellence, effectiveness and accountability (Foote, 1998). "The overriding purpose of evaluation is clearly to improve the teaching program, to move toward faculty excellence" (Seldin, 1980, p. 157). Seldin reiterated in 1984 that there is no greater intent for performance evaluations than to improve performance. Seldin (1980) further asserted

that the professed purpose of the SET at most institutions is improving teaching but that in practice, student evaluations are used administratively. Zelby (1974) contended that SET may adversely affect educational quality if they are used in determination of salary and promotion.

“The cornerstone of every faculty evaluation program is its purpose: the purpose influences the kind of questions asked, sources of data, depth of data analysis, and dissemination of findings” (Seldin, 1984, p. 127). When modern evaluations began, the fundamental purpose was to provide information that could be beneficial to students and faculty (Zelby, 1974). In higher education today, the principal purposes of student evaluations are to improve performance and provide rationale for administrative decisions on tenure, promotion and retention (Seldin, 1984). Other purposes served by evaluations are (a) aiding student choice of courses and instructors, (b) measuring the quality of the course for curriculum development, (c) providing process description for research on teaching, (d) clarifying future performance expectations of the faculty and administrators, and (e) maintaining communication between faculty and administration (Marsh, 1987; McKeachie, 1979).

In a perfect world, faculty evaluations for purposes of improving teaching would be conducted separately from evaluations for personnel decisions (Seldin, 1984). For improving teaching, ratings would provide an accurate diagnosis of problems and solutions (McKeachie, 1979). If the intended use was for personnel decisions, ratings would be valid measures of teaching effectiveness. For student guidance, ratings would contribute valid information allowing students to choose the more valuable learning experience (McKeachie, 1979). Given time and fiscal restraints, it is doubtful that

higher education can accomplish separate ratings (Seldin, 1984). “Societal and professional willingness to pay the cost of objective achievement measures appears to be confined primarily to specialized training programs with important applied foci, such as medicine, law, or architecture” where the purpose of assessment is certification of graduates (Greenwald, 1998).

Formative Student Evaluations of Teaching

In 1984, Seldin wrote:

Faced by an economic squeeze unprecedented in recent years for its severity and duration, the nation’s colleges and universities are struggling to cope with reduced budgets, hunting for new money sources, and casting a gimlet eye on which faculty to promote and which courses to teach. (p. 1)

Seldin also noted that along with the fiscal problems, postsecondary institutions would have to cope with appreciably smaller numbers of college applicants and competition from corporations becoming involved in education ventures of their own.

Out of financial strain came two notable changes for higher education (Simpson & Siquaw, 2000). In the 1990s, the pressure on higher education increased as state governments and the public expressed their dissatisfaction with higher education costs and outcomes by considering performance indicators when budgeting public monies (Simpson & Siquaw, 2000). In 2000, 50% of the states considered higher education’s performance when allocating money with an expected 70% to follow suit within five years (Simpson & Siquaw, 2000). Also, the “student as a customer” approach to education delivery became increasingly popular as a way to guarantee instructor

performance (McCullough & Gremler, 1999). The student oriented approach led to charges that institutions pander to students by watering down curricula and inflating grades to obtain customer satisfaction (Swenson, 1998; Wilson, 1998).

Public scrutiny and criticism of professors made teaching quality a priority for administrators (Frost & Fukami, 1997). Evaluation is identified as a necessary part of teaching and learning (Foote, 1998). Additionally, Johnson and Kelley (1998) asserted that leaner financial times call for scrutiny of faculty evaluation procedures and performance in general. “In an era of growing accountability and outcomes evaluations, achieving a better understanding of the evaluation of teaching effectiveness may be a necessary step toward including scholarship of teaching in decisions on faculty tenure and promotion” (Hobson & Talbot, 2001, p. 34).

In 1976, the Southern Regional Education Board defined the SET process as having formative and summative purposes. Formative evaluations are intended as benign diagnostic tools to support faculty development and self-improvement (Centra, 1994; Hobson & Talbot, 2001; Marsh & Dunkin, 1997; Simpson & Sigquaw, 2000; Stapleton & Murkison, 2001; Trout, 1997; Zelby, 1974). Formative evaluations reflect how one might have done better (Centra, 1994). Student evaluations serve a formative purpose when four conditions are met: instructors must (a) learn something new from them, (b) value the information, (c) understand how to make the suggested improvements, and (d) be motivated to make improvements (Centra, 1993).

Education is now in a period in which administrators increasingly depend on student evaluations to decide tenure and promotion (Wilson, 1998). Despite differences

in opinions and research findings, SET are essentially and almost universally designed as a measure of teaching effectiveness (Hobson & Talbot, 2001; Marsh, 1987).

Under strained economic and social conditions, higher education institutions will be held increasingly accountable for their roles while being expected to be responsive to society for making teaching more effective and efficient (Perry & Smart, 1997). Student evaluations are one measure of teaching effectiveness that can be used to respond to the mounting public demands for teacher accountability and serve as evidence of efforts higher education is making to measure accountability (Seldin, 1980; Simpson & Siquaw, 2000).

Summative Student Evaluations of Teaching

Student evaluations of teaching have assumed greater importance in the last 20 years because of the need for impartial information on which to appraise instructors for reappointment, tenure and promotion (Williams & Ceci, 1997). Using SET to evaluate faculty excuses administrators from judging the quality of instructors (Johnson & Kelley, 1998; Zelby, 1974) and fulfills demands for greater accountability of faculty (Wachtel, 1998; Williams & Ceci, 1997). Student evaluations are prevalent because they are easy to administer and score (Seldin, 1984), produce numbers that appear reliable and are inexpensive (Williams & Ceci, 1997).

Summative evaluations provide a measure of teaching effectiveness as part of the administrative determination for salary and promotion decisions (Marsh & Dunkin, 1997; Stapleton & Murkison, 2001; Trout, 1997). Student evaluations are convenient and provide an ostensibly objective numerical measure (Stapleton & Murkison, 2001; Wilson,

1998). When evaluations are used summatively, it is reasonable to ask only for positive examples of teaching effectiveness (Centra, 1994).

Even though SET ratings are widely accepted, some researchers see problems with the process (Marlin, 1987; Wachtel, 1998). One major cause of the controversy surrounding SET is the formal, quantitative use of results for promotion and tenure decisions (Zelby, 1974). The use of evaluations in a formal sense changes the nature of the SET process from a helpful collection of information to a device that could become detrimental to education (Zelby, 1974). The once innocuous tool for feedback to instructors developed into a surveillance and control tool for administrative decisions (Trout, 1997). Centra (1993) wrote that any use other than formative alters the effects in the teacher and the role of the evaluators.

Recognizing that SET may be permanent in higher education, some question whether the use of student evaluations has the potential to undermine traditional purposes of higher education institutions (Edwards, 2000; Stapleton & Murkison, 2001). Using student evaluations for determining faculty merit makes the instrument vitally important to faculty which may encourage activities designed to affect SET scores rather than improve instruction (Simpson & Siquaw, 2000). Trout (2000) concluded that some academy members believe the use of numerical forms has led to a decline of rigor and standards, but expects that the deeply entrenched practice of student evaluations would be difficult to unseat.

Johnson and Kelley (1998) warned that a main effect of evaluations of performance is short-term thinking which may work against educational objectives. Trout (2000) agreed that the SET process is well-established in higher education

institutions and stressed that some consider numerical forms used to reward and punish as a detriment to teaching principles. Some researchers have noted that SET may contribute to grade inflation if faculty have an incentive to increase their evaluations (Krautman & Sanders, 1999). Others have maintained that there is no evidence that SET improve student learning and question whether higher education would function better in the absence of student ratings (Greenwald, 1998).

Consequences of Student Evaluations of Teaching

The copious research on student evaluations of teaching provides evidence that the evaluations are multidimensional, valid, reliable, unbiased and useful to students, faculty and administration (Centra, 1993; Marsh & Dunkin, 1997; Marsh & Roche, 1997; McKeachie, 1997; Perry, 1997; Simpson & Siquaw, 2000; Spencer & Flyr, 1992). Many confounding factors have been investigated including student characteristics, instructor characteristics, course characteristics and environmental characteristics (D'Apollonio & Abrami, 1997; Read, Rama, & Raghunandan, 2001; Wachtel, 1998; Williams and Ceci, 1997). However, the effects of SET on teaching generally have not been investigated even though the omission was recognized as early as 1975 (Kulik & McKeachie, 1975; Ryan et al., 1980; Spencer & Flyr, 1992; Stratton & Myers, 1994; Wachtel, 1998).

The Effects of Student Evaluations of Teaching on Instructional Practices

Evidence is mixed as to whether the SET process results in improved teaching (Seldin, 1980). Some researchers have found that teachers who tried to strengthen weaknesses identified in SET did receive better evaluations (Wilson, 1998). Gage (1972)

identified specific influences that led to teaching improvement. Some researchers have reported no significant changes as a result of SET information (Centra, 1972; Greenwald & Gilmore, 1997; Miller, 1971). “On balance, however, enough hard evidence is lacking to prove that student evaluations automatically improve classroom teaching” (Seldin, 1980, p. 38). Millman (1982) agreed and stated that few studies have been conducted that investigate the effect of SET on college level instruction.

Zelby (1974) wrote that teaching in order to get higher SET ratings may not be consistent with the best educational practices. Instructors are fallible and incentive driven like everyone else and may incorporate practices designed to increase ratings, especially when ratings are used in the promotion process (Kipps, 1975; McKenzie (1975); Simpson & Siquaw, 2000; Stapleton & Murkison, 2001). “...If the university administration used SET scores to measure teaching effectiveness and reward faculty who receive high scores, one would expect faculty to search out the least-cost method of raising SET scores in their classes” (Stapleton & Murkison, 2001, p. 6). Wallace and Wallace (1998) wrote that one could buy evaluations by decreasing workload, decreasing the difficulty of exams, spoon feeding material directed to the exam and decreasing grading standards, but none of these behavioral outcomes is consistent with the mission of education.

Ryan et al. (1980) concluded that mandatory numerical-evaluation programs affected the rigor of classroom instruction. The most frequently reported changes were for lowered class work demand, including lowered difficulty of course content and lowered grading standards, both of which may be factors in grade inflation. The early warning of the possibility of counterproductive actions as a result of the SET went

unheeded. Only two more studies on the subject can be found in the next 22 years (Simpson & Siquaw, 2000; Spencer & Flyr, 1992).

The Effects of Student Evaluations of Teaching on Instructor Morale

Ryan et al. (1980) found that the greater proportion of respondents believed that the SET process had more negative than positive effects on their own and their colleagues' morale and job satisfaction. Faculty report a lowered commitment to the job and the institution when ratings were less than expected (Simpson & Siquaw, 2000). These results are supported by industrial research that found employees receiving less than outstanding ratings may experience a significant drop in commitment to the organization (Pearce & Porter, 1986). Gabris and Mitchell (1988) found that receiving less than expected ratings may lead to feelings of apprehension about the organization. Other researchers have reported that performance ratings and overall job satisfaction were directly related to future performance (Blau, 1999; Kohli & Jaworski, 1994). Some researchers believe deteriorating quality of faculty work life will ultimately contribute to a decline in the overall quality of postsecondary institutions (Des Jarlais, 1995).

In 1992, Hackman observed that higher education is experiencing a "crisis of confidence" (Des Jarlais, 1995). Three forces that undermine the morale of faculty at many institutions during stressful times are identified by Johnsrud (1996) as reduced resources and restructuring, loss of credibility with the public and increased internal conflict. Some believe the deteriorating quality of faculty work life contributes to a decline in the quality of higher education institutions (Des Jarlais, 1995).

“Research has documented the importance of morale to performance” (Johnsrud, 1996, p.4). People who feel good about their work perform better (Johnsrud, 1996). Further, the vitality of faculty has a direct impact on academia’s mission of teaching, research and service (Des Jarlais, 1995). According to the Faculty Professional Self-Esteem paradigm, when teachers feel respected and valued, they are more effective (Halford, 1994).

Faculty vitality is crucial to the institution because when the professoriate is threatened the entire higher education system is in jeopardy (Altbach, 1987). Raising faculty morale was identified as one of the crucial tasks facing higher education in the 1990s (Budig, 1986; Des Jarlais, 1995; Kerr & Gade, 1987). Presidents of higher education institutions have identified that low faculty morale will be a key issue for some time (Budig, 1986).

“Our morale is our commitment to move forward, our enthusiasm to take on new challenges, and our spirit to maintain the highest of standards and quality” (Johnsrud, 1996, p. 129). The role of the institution in building morale requires institutionally based assessment (Johnsrud, 1996). In addition, more research is needed to gauge faculty member’s views of SET application and its impact on their teaching content and style.

Faculty Opinions of Student Evaluations of Teaching

Student evaluations of teaching are primarily and generally a measure of teaching effectiveness assessing the quality of instructors on chosen instructional dimensions (Hobson & Talbot, 2001; Marsh, 1987; Simpson & Siquaw, 2000). The abundant research on the SET has not included faculty beliefs about evaluations that affect their

opinions of information from evaluations (Ryan, et al., 1980; Spencer & Flyr, 1992; Stratton & Myers, 1994; Wachtel, 1998).

Faculty Opinions of the Validity of Students as Evaluators

In 1928, Remmers wrote that one must consider it important to note if student attitudes towards instructors were influenced by grades. “Students who think they are getting As tend to think more highly of their professor than students who believe they are getting Cs” (Greenwald & Gilmore, 1997, p. 1209). Greenwald and Gilmore (1997) further noted that students infer course quality from received grades and give high ratings in appreciation for lenient grading. McKenzie (1975) found evidence that if the student and teacher agree on grades, content and classroom entertainment that SET ratings are higher. Wallace and Wallace (1998) maintained that student evaluations are a gauge of student happiness at the conclusion of the course before grades are known.

Student evaluations are criticized in the literature for containing items students cannot properly assess and leaving out demographic and background questions that are recognized sources of response bias (Read et al., (2000). Wallace and Wallace (1998) concurred that, as the scope of SET widened, subject matter included in SET questions exceeds what students are qualified to assess. McKeachie (1979) noted that students cannot judge all aspects of teaching equally well. McKeachie found that student ratings are highly valid as indicators of the achievement of attitudinal and motivational educational goals and reasonably valid as indicators of the achievement of cognitive goals. He believed that judgments of appropriateness of content, goals and level of achievement are probably more competently made by peers.

Other researchers argued that students were the best evaluators because they have a unique vantage point to offer commentaries and suggestions to teachers and some basic assurances to administrators (Wallace & Wallace, 1998). “It is manifestly true that the only direct, daily observers of a professor’s classroom teaching performance are the students in the classroom” (Seldin, 1980, p. 36). Marsh (1998) agreed and believed that as higher education is organized today, student raters have spent more time observing the instructor than anyone else. Students as raters of instruction appear to be the obvious and pragmatic choice (Feldman, 1997).

Human resource research into similar relationships shows that only some aspects of the manager’s work are appropriate for evaluation by subordinates (Bernardin, Dahmus, & Redmon, 1993). The most appropriate skills subordinates are qualified to evaluate are the leader’s communication skills, interpersonal relationship skills, and leadership and management skills (McEvoy, 1990). Ratings may be undermined if other aspects are included (Bernardin, Dahmus & Redmon, 1993).

Faculty Opinion about the appropriate Uses of Student Evaluations of Teaching

Faculty appear to be undecided about the appropriateness of using student ratings administratively for tenure and promotion decisions (Rutland, 1990; Ryan et al., 1980; Zoller, 1992). “One of the main obstacles to effective faculty evaluations is the inability to reach consensus on the evaluations intended purpose” (Rifkin, 1995, p. 1). There is long-standing disagreement over whether evaluation systems can be both formative, used for supporting faculty growth and self-improvement, and summative, used to make personnel decisions, and still be effective (Rifkin, 1995). Some researchers question

whether one instrument can serve both formative and summative purposes (Rifkin, 1995; Simpson & Siquaw, 2000).

Numerical evaluation forms may create an incentive for instructors to do the wrong thing by trying to please students instead of teaching them (Trout, 2000). Simpson and Siquaw (2000) stated that ratings are important to faculty for both psychological and economic reasons. Psychologically, student ratings may be malicious to those who actually are trying to teach well. Economically, student ratings may be used as performance criteria that affect promotion and salaries. The extent of these factors may encourage faculty to try to influence SET ratings (Simpson & Siquaw, 2000).

Zelby (1974) stated that the use of evaluations for summative purposes changed the nature of ratings from a helpful collection of information to a device that could become detrimental to education. Zelby (1974) predicted that the SET process could inhibit experimentation and development if used formally for determination of salaries and promotions. Carroll (2002) feared that adjuncts and instructors who are untenured feel they must design the content of courses to please students.

Faculty Opinion of Their Ability to Improve Student Evaluations of Teaching

Ignoring or trying to influence SET information with inappropriate activities may be destructive to educational objectives (Simpson & Siquaw, 2000). Some instructors design instructional changes to increase SET ratings (Simpson & Siquaw, 2000). Changes have included grade inflation, dumbing down of coursework and lowered workload which may be counter to educational objectives (Ryan et al, 1980; Zelby,

1974). Typical evaluation approaches cannot differentiate between changes made to enhance learning and changes made to increase student ratings (Stapleton & Murkison, 2001; Stratton & Myers, 1994). Several researchers reported that the majority of faculty believed student ratings were influenced more by irrelevant personal qualities of the faculty than by academically relevant activities (Holden 1997; Nesbit & Wilson, 1977; Ryan et al, 1980; Simpson & Siquaw, 2000). The effects of personality are consistent across other professions lending credence to higher education research (DeCarlo & Leigh, 1996). A substantial proportion of respondents did not believe they could improve SET ratings and were not likely to try to do so (Ryan et al, 1980).

Faculty Use of Student Evaluations of Teaching

Faculty beliefs about SET information range from reliable, valid and useful to unreliable, invalid and useless (Aleamoni, 1981). Most researchers today believe SET are valid, reliable and worthwhile (Centra, 1993; McKeachie, 1990; Seldin, 1993). However, the anecdotal evidence of faculty opposition towards the evaluations is extensive (Franklin & Theall, 1989). Simpson and Siquaw (2000) found that instructors perceive problems with student ratings of teaching and suggested more research on the pervasiveness of disenchantment with the SET process and what changes in instruction were made in response to SET ratings.

Ryan et al. (1980) found that 38% of respondents were not inclined to modify teaching to earn higher SET scores and another 30% were only slightly inclined to make changes. Half of the respondents did not believe they could improve SET ratings by any amount through academically relevant activities. Forty percent of respondents believed

they could increase their student ratings through academically irrelevant activities.

Spencer and Flyr (1992) found that 77% of responses indicated student evaluations were not taken into account and 73% reported student ratings never or only occasionally lead to instructional improvement.

Franklin and Theall (1989) found that faculty familiar with SET literature made better use of their ratings. Positive attitudes should be associated with better practices and those ignorant about SET ratings were at greater risk of using bad practices. Generally, faculty appeared not to be aware of current literature that might help them make informed decisions about how to use SET information.

“Whenever an individual chooses between alternatives which involve uncertain outcomes, it seems clear that his behavior is affected not only by his preference about these outcomes, but also by the degree to which he believes these outcomes to be probable” (Vroom, 1964, p. 17). Expectancy Theory offers an explanation of the effects of faculty opinions of their ability to improve SET ratings on changes in instructional practices and morale. Faculty appeared not to be aware of current literature that might help them make informed decisions about appropriate uses of the SET. That unawareness may lead instructors to expect that the SET will have harmful impact. Under Expectancy Theory, such negative expectations prevent faculty from viewing SET favorably and gaining the best possible benefit from it.

Expectancy Theory

Vroom is credited with developing Expectancy Theory in the 1960s. The premise serves as the foundation for a number of other theories sharing this name. Expectancy Theory is a process theory, dealing with how motivation occurs rather than what motivates people, and hypothesizes that motivation and effort to perform is a function of the probability that the required behavior will result in the desired outcome (Vroom, 1964). Alternatively stated, motivation depends on how much we want something and our likelihood of getting it. Behavior is selected from alternatives that maximize pleasure, minimize pain and produce the best rewards. The focus of the theory is the individual's motivation to behave in a certain way in a particular situation rather than an individual's overall motivation.

When choosing between behaviors, individuals select the choice with the highest motivational force. Motivation energizes, guides, and sustains behavior. Motivational force is a product of the three perceptions of expectancy, instrumentality and valence. If any one of the perceptions is low, motivation and effort is low.

Expectancy Theory addresses the strength of one's belief that the effort put forth will result in successful performance. What is the perceived likelihood that effort will lead to task accomplishment? Deciding whether the outcome is possible is influenced by previous successes, self confidence and individual skill. Individuals will attempt a task only if they believe that it can be done (Top Education, 2001; Vroom, 1964).

Instrumentality is one's perceived probability that a high level of performance will result in desired outcomes. That is, performance will be instrumental in leading to particular outcomes. If one does meet performance expectations, he or she expects to

receive a greater reward perhaps through a pay increase, promotion, recognition or a sense of personal accomplishment. Individuals must understand the connection between desired outcomes and performance for instrumentality to be high (Top Education, 2001; Vroom, 1964).

Valence is subjective and refers to the value of the expected rewards to the individual or the level of satisfaction the individual expects from the outcome. Valence is a function of needs, goals, values and preferences (Top Education, 2001; Vroom, 1964).

Individuals are motivated to perform if they perceive that the effort exerted will lead to the desired performance and the given performance will lead to desired outcomes. These two perceptions of expectancy and instrumentality interact with each other and with the valence, or value, of outcomes to ascertain the general level of motivation. People select the level of performance that will best meet their needs, goals and values (Top Education, 2001; Vroom, 1964).

For this study, expectancy relates to whether instructors believe their efforts will lead to improved student learning and/or promotion, tenure and raises. Instrumentality is whether instructors believe that improved student achievement will earn them a boost in morale, tenure, promotion or raises. Valence refers to whether individual instructors value the positive outcomes of the satisfaction of seeing students succeed or personal career rewards. Valence can include whether the individual instructor wants to prevent harmful consequences related to not meeting student achievement goals (Odden, 2000).

The motivation to improve teaching to increase student ratings of instruction depends on the importance of outcomes and the belief as to whether higher ratings will

result in desired outcomes such as promotion, tenure and raises. According to Expectancy Theory, SET process is more important to faculty when evaluations are used in considerations concerning promotions or used to improve student learning. If faculty value promotions or increased student learning, faculty will be expected to exert substantial effort to increase ratings of their classes. The question becomes what behavior do instructors change in order to influence SET ratings? Will student demands be satisfied by better teaching or undeserved higher grades?

Motivation will be low if individuals do not believe they can be successful at the task, or if they believe that the successful task will not lead to positive outcomes, or if they believe outcomes will be negative. If instructors believe that receiving higher SET ratings is too difficult or that they cannot influence the ratings, student comments may be ignored and morale may be low. Also, if instructors believe the goals of promotion and tenure are not possible, morale may be low.

CHAPTER III

Methods

This study investigated instructors' opinions of perceived effects of student evaluations of teaching (SET) on instructional practices and instructor morale. Specifically, this investigation collected information about faculty opinion of the validity of students as evaluators, faculty opinion about the appropriate uses of SET information, faculty opinion about their ability to improve SET ratings and faculty interest in SET ratings. The following questions were investigated in this study:

1. Is there a statistically significant relationship between faculty opinion of the validity of student evaluations of teaching and faculty perceptions of changes in instructional practices?
2. Is there a statistically significant relationship between faculty opinion of the validity of student evaluations of teaching and faculty perceptions of changes in morale?
3. Is there a statistically significant relationship between faculty opinion of the appropriate uses of student evaluations of teaching and faculty perceptions of changes in instructional practices?
4. Is there a statistically significant relationship between faculty opinion of appropriate uses of student evaluations of teaching and faculty perceptions of changes in morale?
5. Is there a statistically significant relationship between faculty perceptions of their ability to influence ratings from student evaluations of teaching and faculty perceptions of changes in instructional practices?

6. Is there a statistically significant relationship between faculty perceptions of their ability to increase ratings from student evaluations of teaching and faculty perceptions of changes in morale?
7. Is there a statistically significant relationship between faculty interest in student evaluations of teaching and faculty perceptions of changes in instructional practices?
8. Is there a statistically significant relationship between faculty interest in student evaluations of teaching and faculty perceptions of changes in morale?

Population and Sample

The population was the faculty members of public, four-year colleges in West Virginia during the 2002-2003 academic year. The institutions were Bluefield State College, Concord College, Fairmont State College, Glenville State College, Shepherd College, West Liberty State College, West Virginia State College and West Virginia University Institute of Technology. All eight institutions were in the same Carnegie Classification which groups institutions by commonalities (Carnegiefoundation.org/Classification/CIHE2000/background.html). All were listed as *Liberal Arts Colleges* in 1973 and changed to *Baccalaureate Colleges* in 1994. The particular classification was chosen from all higher education classifications because the faculty of Liberal Arts/Baccalaureate Institutions reported the highest interest in teaching (Carnegie Foundation for the Advancement of Teaching, 1989 National Survey). Historically, without the demands for research and publication, the classification of

institutions has maintained the scholarship of teaching as the central mission (Boyer, 1990). Student evaluations of teaching were best tested where teaching is most valued.

Lists of faculty members (N= approximately 935) were solicited from the chief academic or administrative officer from each institution. An appropriately sized random sample (n= approximately 274) was chosen from this population to participate in the study (Johnson & Christensen, 2000). The return rate of 70.7% exceeded the recommended minimum level of 50% plus one which strengthened the results of the study and supported the generalization of findings (Kerlinger & Lee, 2000).

Instrumentation

Participants were mailed the *Gall Faculty Response to Evaluations of Teaching (FRET) Survey*, a self-report questionnaire, which gathered information related to faculty opinions about the effects of the SET process on instructional practices and instructor morale (see Appendix A). Two items related to the instructor's opinion of the validity of students as evaluators. Faculty opinion about the appropriate uses of SET information was surveyed through two items concerning evaluations as an aid to teaching and evaluations used administratively for retention, promotion and tenure. Faculty opinion about their ability to improve SET ratings was examined through two items. Faculty use of SET information was measured through two items soliciting information related to the individual instructor's personal use of voluntary evaluations and mandatory evaluations. The survey also included demographic information, a practice used by many researchers to investigate how attitudes and behavior differ for people with various attributes (Dillman, 1978).

The *Gall FRET Survey* was inspired by research literature, especially the work of Ryan et al. (1980), and developed by the researcher. The instrument was pilot tested with ten people similar to those in the study in an attempt to establish reliability and validity (Johnson & Christensen, 2000). Pilot testing also helped establish the readability of the instrument and determine how long it takes to complete the instrument (Johnson & Christensen, 2000). It was not necessary to revise and pilot test the instrument a second time.

Methods

Self-report questionnaire survey procedures were used to investigate faculty perceptions of the effects of student evaluations of teaching on instructional practices and instructor morale (Campbell & Staley, 1963; Kerlinger & Lee, 2000). This study provided for participant anonymity to reduce the effects of response bias (Kerlinger & Lee, 2000).

A packet containing a cover letter (see Appendix B), *Gall Faculty Response to Evaluations of Teaching Survey*, and a self-addressed stamped return envelope was mailed to each person in the survey sample population. The cover letter explained the purpose of the study, assured anonymity of subjects and encouraged participation in the study. Subjects were asked to respond to the survey and return them within one week. One week after the initial mailing, a reminder postcard (see Appendix C) was mailed to encourage completion and return of the survey. A follow-up packet containing a follow-up cover letter (see Appendix D), survey and self-addressed stamped return envelope was mailed two weeks after the initial mailing.

Data Analysis

Descriptive statistics were used to explain and summarize the quantitative data collected in a more concise and convenient form (Kerlinger & Lee, 2000). Descriptive statistics help communicate and interpret the data. Measures of central tendency are single numerical values chosen to be typical of the collection. Measures of dispersion are single numerical values that reveal information about the manner in which data are distributed. Frequency distributions assist in the interpretation of a collection of data by arranging measures of a given variable to indicate the frequency of occurrence of the different values. This allowed the researcher to determine at a glance the general distribution of the data.

Correlation analysis refers to the relationship between two variables or the degree to which two variables are related and follow the same linear path (Kerlinger & Lee, 2000). The purposes are to learn about the relationship among variables and to make predictions based on the relationships. Post hoc analyses were conducted where appropriate.

Faculty opinion of the validity of students as evaluators, faculty opinion about the appropriate uses of SET information, faculty opinion about their ability to improve SET ratings and faculty use of SET information were the independent variables (Johnson & Christensen, 2000). Instructional practices and instructor morale were the dependent variables.

Summary

The procedures described in this chapter were designed to determine the effects of faculty opinion of the validity of students as evaluators, faculty opinion about the appropriate uses of SET information, faculty opinion about their ability to improve SET ratings and faculty use of SET information on instructional practices and instructor morale. A random sample of faculty members from public, four-year colleges in West Virginia was surveyed. Appropriate statistical tests were performed to answer each of the research questions as posed.

CHAPTER IV

Analysis of Data

Student evaluations of teaching (SET) are the most used instrument for evaluating teaching in higher education. The composition of and reasoning for using the SET process has been well researched but the relationship between SET and instructional practices and instructor morale has been investigated by only a few researchers. This study investigated faculty perceptions of the effects of student evaluations of teaching on instructional practices and instructor morale.

The population (N= approximately 935) was the faculty of the eight public four year colleges in West Virginia from which a random sample was chosen to receive the questionnaire. An appropriately sized random sample (n= approximately 274) was chosen to participate in the study. Fourteen faculty were eliminated because of outdated addresses. The final working sample was 260. One-hundred eighty four questionnaires were returned, for a return rate of 70.7%.

The independent variables were faculty opinion of the validity of students as evaluators, faculty opinion about the appropriate uses of SET, faculty opinion of their ability to improve SET ratings, and faculty use of SET information. The dependent variables were the perceived effects of SET on instructional practices and instructor morale. Demographic data of sex, faculty status, college/school, ethnicity, age, years in higher education, and years at present institution were also requested.

Statistical Analysis of the Data

Data were gathered using the researcher-developed Gall *Faculty Response to Evaluation of Teaching Survey (FRET)*. The *FRET* asked 30 questions of which eight were demographic and 22 were generated by research literature. A four-point Likert scale for “strongly agree” to “strongly disagree” was used for all research questions.

The four independent variables were measured by two questions each. The mean of each set of two questions was calculated to correlate with the means of the dependent variable responses.

The dependent variables were measured by several questions. The perceived effects of SET on instructional practices were measured by 12 questions asking respondents to rate how much the information from the SET had encouraged them to make changes in their instructional routines. A mean of the questions was calculated to correlate with the means of the independent variables.

The second dependent variable, the perceived effects of the SET process on instructor morale, was measured by three questions concerning faculty self-images and job satisfaction. A mean was calculated to correlate with the independent variables. Results of the correlations between the independent and dependent variables are presented in Table 1.

Table 1

Faculty Perceptions of the Relationship of Student Evaluations of Teaching on Instructional Changes and Instructor Morale

Faculty Opinion of	Correlation Coefficient for Changes in	
	Instructional Practices	Morale
Students as Evaluators	.409**	-.065
Appropriate Uses of SET	.459**	.083
Ability to Influence Ratings	.301**	.089
Interest in SET	.520**	.057

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

Research Question 1: Is there a statistically significant relationship between faculty opinion of the validity of students as evaluators of teaching and faculty perceptions of changes in instructional practices?

Questions 8 and 9 on the questionnaire asked respondents to rate their opinion of whether students were capable of evaluating teaching and whether mandated SET process was an appropriate method of evaluation. A mean of the questions was calculated and correlated to the mean for perceived changes in instructional practices. The correlation coefficient for the two means was .411, significant at the alpha 0.01 level. Therefore,

there is a significant relationship between faculty opinion of the validity of students as evaluators of teaching and faculty perceptions of changes in instructional practices.

Research Question 2: Is there a statistically significant relationship between faculty opinion of the validity of students as evaluators of teaching and faculty perceptions of changes in morale?

The mean for questions 8 and 9 was correlated to the mean for perceived effects on instructor morale. The result was not significant at the alpha level of 0.01. Therefore, there is no significant relationship between faculty opinion of the validity of students as evaluators of teaching and faculty perceptions of changes in morale. This was the only negative correlation in the study.

Research Question 3: Is there a statistically significant relationship between faculty opinion of the appropriate uses of student evaluations of teaching and faculty perceptions of changes in instructional practices?

Questions 10 and 11 on the questionnaire asked faculty to rate their opinion of the propriety of using SET results for improving teaching and/or administrative decisions. The mean of these questions was correlated to the mean for the perceptions of changes in instruction and had a coefficient of .451, significant at the alpha level of 0.01. Therefore, there is a significant relationship between faculty opinion of the appropriate uses of student evaluations of teaching and faculty perceptions of changes in instructional practices.

Research Question 4: Is there a statistically significant relationship between faculty opinion of appropriate uses of student evaluations of teaching and faculty perceptions of changes in morale?

The mean of questions 10 and 11 on the questionnaire was correlated to the mean for the perceived effects on instructor morale. The result was not significant at the alpha 0.01 level. Therefore, there is not a significant relationship between faculty opinion of appropriate uses of student evaluations of teaching and faculty perceptions of changes in morale?

Research Question 5: Is there a statistically significant relationship between faculty perceptions of their ability to influence ratings from student evaluations of teaching and faculty perceptions of changes in instructional practices?

Questions 12 and 13 on the questionnaire asked faculty to rate whether professors thought they could improve SET ratings through instructional changes or other enjoyable activities. The mean of the two questions and the mean of the perceived effects on instructional practices had a correlation coefficient of .297 which is significant at the alpha 0.01 level. Therefore, there is significant relationship between faculty perceptions of their ability to influence ratings from student evaluations of teaching and faculty perceptions of changes in instructional practices.

Research Question 6: Is there a statistically significant relationship between faculty perceptions of their ability to increase ratings from student evaluations of teaching and faculty perceptions of changes in morale?

The mean of question 12 and 13 on the questionnaire was correlated to the mean for perceived effects on morale. The correlation was not significant at the alpha 0.01 level. Therefore, there is not a significant relationship between faculty perceptions of their ability to increase ratings from student evaluations of teaching and faculty perceptions of changes in morale.

Research Question 7: Is there a statistically significant relationship between faculty interest in student evaluations of teaching and faculty perceptions of changes in instructional practices?

Questions 29 and 30 on the questionnaire asked faculty to rate how much they used SET information. The mean from these questions correlated to the mean of perceived changes in instructional practices with a coefficient of .508, significant at the 0.01 alpha level. Therefore, there is a significant relationship between faculty interest in student evaluations of teaching and faculty perceptions of changes in instructional practices.

Research Question 8: Is there a statistically significant relationship between faculty interest in student evaluations of teaching and faculty perceptions of changes in morale?

The mean of question 29 and 30 on the questionnaire was correlated to the mean of perceived effects on morale and was not significant at the 0.01 alpha level. Therefore, there is a significant relationship between faculty interest in student evaluations of teaching and faculty perceptions of changes in morale.

Descriptive Data

Of the 184 respondents, 90 (48.9%) reported their sex as female and 94 (51.1%) reported their sex as male. The frequency distribution for the faculty status of respondents in this study is reported in Table 2.

Table 2

Frequency Distribution by Faculty Status

	Frequency	Percent
Tenure Track- Tenured	108	58.7
Tenure Track- Not Tenured	29	15.8
Not Tenure Track	16	8.7
Adjunct/Part Time	29	15.8
Unreported	2	1.1
Total	184	100.0

The frequency distribution for the college/school of respondents is reported in Table 3.

Table 3

Frequency Distribution for College/School

	Frequency	Percent
Applied Sciences	45	24.5
Education	23	12.5
Business	24	13.0
Humanities	31	16.8
Social Sciences	21	11.4
Fine Arts	18	9.8
Other	17	9.2
No Response	5	2.7
Total	184	100.0

The Frequency distribution for the ethnicity of respondents is reported in Table 4.

Table 4

Frequency Distribution for Ethnicity

	Frequency	Percent
American Indian/Alaskan Native	2	1.1
Asian	1	0.5
Native Hawaiian/Other Pacific Islander	0	0.0
Black/African American	7	3.8
Hispanic/Latino	0	0.0
White	170	92.4
Other	2	1.1
No Response	2	1.1
Total	184	100.0

The remaining descriptive data are found in the final tables. The frequency distributions for years in higher education is reported in Table 5, age in Table 6, and years at present institution in Table 7.

Table 5

Frequency Distribution for Years in Higher Education

	Frequency	Percent
0- 5 years	35	19.0
6-10 years	23	12.5
11-15 years	37	22.2
16-20 years	20	10.8
More than 20 years	68	40.0
No Response	1	0.5
Total	184	100.0

Table 6

Frequency Distribution for Age

Years	Frequency	Percent
27-35	18	9.6
36-45	32	17.4
46-55	65	35.3
56-65	25	13.6
66 and older	7	2.6
Missing	3	1.6

Total	184	100.0
-------	-----	-------

Table 7

Frequency Distribution for Years at Present Institution

Years	Frequency	Percent
1-5	58	31.6
5-10	34	18.4
11 or more	92	50.0
Total	184	100.0

Ancillary Findings

Significant correlations between demographic information and the study's variables were found in two instances. Years at present institution and faculty perception of changes in instructional practices had a correlation coefficient of .221 which is significant at the alpha 0.01 level. Years of higher education teaching experience and faculty perception of changes in instructional practices had a correlation coefficient of .192 which is significant at the alpha 0.01 level. Therefore, the number of years spent working at the same institution and the number of years spent teaching in higher education are significantly related to faculty perceptions of changes in instructional practices.

Summary

Of the 260 faculty in the working sample, 70.7% (n=184) responded to the FRET survey. By a small percentage, most respondents were male (51.1%). Most (58.7%) were tenured or on tenure track. Applied Sciences was the most reported college or school at 24.5%. The overwhelming reported ethnicity was White (92.4%). Forty percent had worked in higher education for more than 20 years. The most frequent age range was 46-55 (35.3%). Exactly half reported they had been at their present institution for eleven or more years.

The independent variables were faculty opinion of the validity of students as evaluators, faculty opinion about the appropriate uses of SET, faculty opinion of their ability to improve SET ratings, and faculty use of SET information. All dependent variables significantly correlated to the faculty perception of changes in instructional practices. None of the dependent variables, however, significantly correlated to faculty perceptions of changes in morale.

Post hoc analysis of demographic data showed significant correlation in two areas. Faculty perception of changes in instructional practices significantly correlated to both years teaching at the present institution and years of higher education teaching experience.

CHAPTER V

Findings, Conclusions and Recommendations

Student evaluations of teaching (SET) are the most commonly used instrument for evaluating higher education faculty (Seldin, 1984; Spencer & Flyr, 1992). Extensive research exists on the composition and reasoning for using the SET (Read, Rama, & Raghunandan, 2001; Wachtel, 1998; Wallace & Wallace, 1998) including research on validity and reliability (Aleamoni, 1987; Cashin, 1988; Marsh, 1984), characteristics of teachers, courses, students and teaching environments (Calderon & Green, 1997; Martin, 1998). Little is known, however, about the potential impact of the SET process on teaching practices and instructor morale (Ryan, Andersen, & Birchler, 1980; Simpson & Siquaw, 2000; Spencer & Flyr, 1992; Stratton & Myers, 1994; Wachtel, 1998). This study investigated faculty perceptions of the relationship between student evaluations of teaching and changes in instructional practices and instructor morale.

Methods

The population for this study was the faculty members of the eight public four-year colleges in West Virginia during the 2002-2003 academic year. All of these were in the same Carnegie Classification which maintains the scholarship of teaching as the central mission (Boyer, 1990). A random sample of 274 faculty was chosen from the population.

Self-reporting questionnaire survey procedures were followed (Campbell & Stanley, 1963; Kerlinger & Lee, 2000). The first mailing included a cover letter, the *Gall FRET Survey*, and a self-addressed stamped return envelope. One week after the

initial mailing, a reminder postcard was mailed. A packet containing a follow-up cover letter, survey and self-addressed stamped envelope was mailed two weeks after the initial mailing.

The mailings were sent to faculty members in the sample population after the start of the fall semester, a relatively slow period in the academic calendar. A total of 184 questionnaires were returned for a rate of 70.7%.

Instrumentation and Data Analysis

The researcher developed *Gall Faculty Response of Evaluations of Teaching (FRET) Survey* was inspired by research literature. This self-reporting questionnaire solicited information about faculty opinions of the relationship of the SET process with perceived changes in instructional practices and instructor morale. Demographic information identified by SET research literature as being related was also requested.

Data was analyzed using *SPSS version 11.5*. Particularly, correlation analyses were used to determine significant relationships among variables. Descriptive statistics were used to summarize demographic data.

Descriptive Data

Of the 184 respondents, 90 (48.9%) were female and 94 (51.1%) were male. The largest faculty status category reported was tenure track-tenured with 58.7% of the respondents falling into this group. Most respondents were teaching in the Applied Sciences (24.5%) with Humanities second at 16.8%. Other college/school teaching assignments percentages ranged from 9.2% to 13%.

The large majority of the respondents indicated their ethnic group to be white (92.4%) which is proportionate with the population demographics of West Virginia. The mean for years of higher education teaching experience was 17.8 with those working more than 20 years comprising the largest category (40.0%). The mean for age was 51 and the largest group was the 46-55 years range (35.3%). The mean for years working at their present institution was 14 with 50% of the faculty reporting 11 or more years.

Findings and Conclusions

Results for research questions investigating the relationship between the four independent variables and faculty perception of changes in instructional practices were significant in all cases (questions 1, 3, 5, and 7). Findings for research questions exploring the relationship between the four independent variables and faculty perception of changes in instructor morale were not significant (questions 2, 4, 6, and 8).

Research Question 1: Is there a statistically significant relationship between faculty opinion of the validity of students as evaluators and faculty perceptions of changes in instructional practices?

A significant relationship was found between faculty opinion of the validity of students as evaluators of teaching and faculty perceptions of changes in instructional practices. The finding indicates that faculty regard students as legitimate evaluators of teaching and use the information students offer. Faculty perceive the information from

student evaluations as potentially useful when considering modifications to their teaching techniques.

This finding is compatible with those of the few other studies on the subject. Some believe students are the best evaluators because they are the only direct, daily observers of classroom instruction (Seldin, 1980). Marsh (1998) argued that students spend more time observing instruction than others in higher education and are, therefore, the appropriate evaluators of instruction. Students are the obvious and pragmatic choice for evaluating teaching.

Research Question 2: Is there a statistically significant relationship between faculty opinion of the validity of students as evaluators of teaching and faculty perceptions of changes in morale?

A statistically significant relationship was not found between faculty opinion of the validity of students as evaluators and faculty perceptions of changes in morale. Faculty appear to appreciate students opinions of their teaching methods but student opinions do not affect faculty self-worth.

Data on the relationship between the SET process and instructor morale is sparse. Ryan et al. (1980) reported a significant reduction in staff morale and diminished job satisfaction as a result of the SET process. Industrial research supports the finding that employees receiving lower than expected ratings experience a significant drop in organizational commitment (Pearce & Porter, 1986). Other researchers reported a direct relationship between performance ratings, overall job satisfaction and future performance (Blau, 1999; Kohli & Jaworski, 1994).

The finding of this study conflicts with research literature on faculty morale, although authoritative conclusions cannot be made with the little empirical data available. Most faculty today started teaching knowing that student evaluations were a part of academic life. Further, faculty in this study are from colleges with the mission of teaching. They should feel that their instructional practices are worthy and accept that their institution's mission requires critical review of teaching strategies with no reflection on personal worth.

Research Question 3: Is there a statistically significant relationship between faculty opinion of the appropriate uses of student evaluations of teaching and faculty perceptions of changes in instructional practices?

A statistically significant relationship was found between faculty opinion of the appropriate uses of student evaluations of teaching and faculty perceptions of changes in instructional practices. The finding indicates that faculty are comfortable with, or at least accepting of, student evaluations being used for both formative and summative purposes. That is, faculty perceive that information from student evaluations is properly used for both professional development and personnel decisions. They perceive that making changes in teaching practices can lead to promotion and tenure.

Authors of previous studies reported that faculty were undecided about the propriety of using SET information for formative and summative purposes (Rutland, 1990; Ryan et al., 1980; Zoller, 1992). As early as 1974, Zelby warned that imprudent use of the SET process would inhibit progressive development of education especially if used to determine promotions. In 2000, Simpson and Siquaw continued the argument

that administrative use of SET results encourages faculty to make changes in teaching methods that are detrimental to student learning. The faculty respondents in this study, however, seemed accepting of the use of student evaluations as a guide to professional development and as part of their administrative evaluation. The SET is institutionalized now and its uses and motives are not as suspect as when the process was new.

Research Question 4: Is there a statistically significant relationship between faculty opinion of appropriate uses of student evaluations of teaching and faculty perceptions of changes in morale?

A statistically significant relationship was not found between faculty opinion of the appropriate uses of student evaluations of teaching and faculty perceptions of changes in morale. Faculty in this study were accepting of the practice of using student ratings formatively and summatively but had isolated morale from the ratings.

The finding for this research question differs from findings from other published research. As the uses of the SET process expanded to include administrative functions, some researchers maintained that subject matter was included that exceeded what students were qualified to rate (Wallace & Wallace, 1998), a finding supported by human resource researchers (Bernardin, Dahmus & Redmon, 1993). Surveys asking for judgments beyond the scope of student understanding undermine confidence in the ratings. Others reported the only appropriate use for formative evaluations is as an aid to improving instruction and that summative use changes the effects on the instructor (Centra, 1993).

As stated previously, faculty in this study have been subjected to SET ratings for most of their professional lives and accept them as a fact of employment. One would

expect teaching acumen to be part of promotion decisions in higher education institutions and respondents in this study report no connection between the use of student ratings and morale.

Research Question 5: Is there a statistically significant relationship between faculty perceptions of their ability to influence ratings from student evaluations of teaching and faculty perceptions of changes in instructional practices?

A significant relationship was found between faculty perceptions of their ability to influence ratings from student evaluations of teaching and faculty perceptions of changes in instructional practices. The finding indicates that faculty believe they can impact SET ratings through altering instructional practices and interacting more with students.

Previous research indicated that faculty believe SET ratings can be most directly influenced through irrelevant personal characteristics (Ryan et al., 1980). Other respondents in the same study reported that ratings could not be improved through academic measures. This does not seem to be the case with faculty respondents in the current study who indicated that modifying classroom strategies, among other things, should affect student ratings. The current push for improved teaching outcomes could have influenced faculty opinion. Again, the SET process was in place when the faculty began teaching and they may view student ratings of their teaching as a given. Whatever the case, the disenchantment with the process and its inherent problems found in research by Simpson and Siquaw (2000) is not evident from the data in this study. Insufficient research literature limits definitive conclusions.

Research Question 6: Is there a statistically significant relationship between faculty perceptions of their ability to increase ratings from student evaluations of teaching and faculty perceptions of changes in morale?

A significant relationship was not found between faculty perceptions of their ability to influence ratings from student evaluations of teaching and faculty perceptions of changes in morale. Faculty appear to disassociate self-esteem from the ability, or lack of ability, to affect changes in ratings.

One would expect confidence to be affected if faculty believed they could not increase student evaluations by better instruction. Further, if factors that did affect ratings could not be altered, morale would be expected to drop (Ryan et al., 1980). Yet the finding in the current study oppose those from the few available studies (Ryan et al., 1980; Simpson & Siquaw, 2000). Faculty might have learned to insulate their self-esteem from student opinion because they receive support from the administrators of their colleges. Continued use of the process could have made faculty cynical, in which case the encouragement and satisfaction from good evaluations would have also been lost.

Research Question 7: Is there a statistically significant relationship between faculty interest in student evaluations of teaching and faculty perceptions of changes in instructional practices?

A significant relationship was found between faculty interest in student evaluations of teaching and faculty perceptions of changes in instructional practices. Respondents indicated that they do consider evaluations of their teaching and that those evaluations impact adjustments to teaching approaches.

Research has established that student evaluations of teaching are valid, reliable, and worthwhile (Centra, 1993; McKeachie, 1990; Seldin, 1993) but some have questioned whether faculty are familiar with those findings (Franklin & Theall, 1989). Positive attitudes toward the student evaluation process should be associated with better practice. Unawareness carries a greater risk of poorer practices and should lead instructors to regard the process as harmful (Franklin & Theall, 1989). Faculty in this study indicated that they do consider student evaluations and suggestions when revising instructional practices. They might be familiar with SET literature and are more accepting of student suggestions. Or, conceivably they have succumbed to pressures to adapt to student opinion and are responding to new administrative attitudes supporting students' right to have input into their education. In either case, whether learning is increased or decreased through faculty knowledge of SET literature and accepting student suggestions is yet to be determined.

Research Question 8: Is there a statistically significant relationship between faculty interest in student evaluations of teaching and faculty perceptions of changes in morale?

A significant relationship was not found between faculty interest in student evaluations of teaching and faculty perceptions of changes in morale. Faculty report that they review information from student evaluations for professional growth but evaluations do not influence their self-confidence.

Halford (1994) found that factors affecting morale are crucial because teachers are more effective when they feel valued and respected. Attention to these issues contributes to the institution's energy (Kang & Miller; 2000).

Understanding the evaluation process and its inclusion in administrative decisions should help shape faculty response to student ratings (Hobson & Talbot, 2001). Possibly faculty in this study have become informed about the SET process and can separate their thoughts about professional skills from thoughts affecting morale. If SET ratings do measure student happiness or agreement with grades (McKenzie, 1975; Wallace & Wallace, 1998), faculty would be well served by disassociating their morale from ratings. However, as discussed above, this protective mechanism can also isolate faculty from the positive effects of favorable evaluations. Research needs to be expanded to establish the mechanisms produced by the SET process.

Ancillary Findings

The respondents' reported years of experience in higher education and years working at their present institution was significantly related to the perception of changes in instructional practices. Experience in teaching and experience with the SET process over time seemed to encourage consideration of student input when teaching approaches were revised. Goldberg and Callahan (1991) reported that new instructors and adjuncts should be particularly sensitive to student input about their teaching abilities, but this is not supported by the results of this study. The increased use of part-time faculty as a cost-cutting measure appears to have resulted in transient faculty who lack commitment to higher education and are, therefore, not as concerned with student opinions of their teaching. Some part-time faculty choose to avoid controversy by awarding better grades.

Goldberg and Callahan (1991) concluded that adjunct faculty tend to grade higher and receive better ratings than full-time faculty which lends credence to accusations of grade inflation. Because their employment is temporary, these instructors may not become involved with institutional practices and development programs. Full-time faculty have more invested in their careers, the process of student learning, and evaluations and they may feel that SET information deserves thoughtful consideration.

Implications for Higher Education Administrators

Because student evaluators are accepted as valid by faculty as confirmed in this study, education and learning should be better served by further and thoughtful use of student opinions. Administrators should provide financial and emotional support for increased use of evaluation and development programs. Gathering SET information for every class instead of selected classes adds to the pool of suggestions promoting faculty personal and professional development. Adding mid-term evaluations would allow faculty to make potentially helpful changes quickly. If the suggestions are not productive, end of term evaluations by the same students should recommend elimination of the poorer practices. Particularly interested faculty should try new techniques and gather student reactions throughout the course using evaluation questions designed for this purpose.

Educating students about the uses of faculty evaluations would bring a new appreciation of the weight of their opinions. Students would give more thought to evaluations if they realize that their suggestions were being seriously considered. A

Socratic dialogue with regard to pedagogy should benefit everyone involved in higher education.

Faculty respondents appear secure using SET information for both formative and summative purposes. Conscientious teachers should welcome the inclusion of evaluations of their teaching in administrative decisions. Using student evaluations summatively provides administrators information directly from those who are being educated thus providing a basis for informed administrative decisions. One problem, however, is that typical evaluations cannot discriminate between student reactions to more effective instruction and reactions to lowered class requirements that diminish learning. Faculty development programs, as discussed later, should deter unconsidered changes leading to grade inflation and dumbing down of coursework.

Student evaluations of teaching are not a bad concept, albeit a poor tool that attempts to serve two different purposes. Empirical evidence from this study indicating faculty acceptance of the dual use of ratings should not eliminate consideration of whether information from student evaluations would be more valuable if separate, more in-depth, assessments were used.

Faculty appear to value student evaluations of their teaching as an important, or at least obligatory, part of their professional development. An instrument to evaluate teaching practices should diagnose instructional problems and provide valuable on task suggestions for improved learning. Faculty should create other opportunities for evaluation by asking students to rate a particular assignment or project. Formative evaluations could be communicated between students and instructors only. Use of an anonymous list of departmental ratings would allow instructors to consider their own

work against others on the same team. A second instrument should be used summatively for determinations of promotions and tenure. Certainly, further delineation and consideration of formative and summative tasks would well serve both faculty and administrators of higher education.

Faculty respondents perceived that making changes in instructional practices should raise their SET ratings. Expectancy Theory hypothesizes that the motivation and effort to perform is a function of the probability that the necessary behavior will result in the desired outcome (Vroom, 1964). Therefore, if faculty believe that changes in teaching strategies should lead to personal satisfaction and/or promotion and raises, Expectancy Theory predicts they would make substantive effort to improve their teaching, possibly justifying administrative use of SET information. Administrators who encourage and reward the scholarship of teaching and faculty commitment to education will strengthen the likelihood of renewed support from the public and government.

Faculty in this study indicated that they did consider suggestions from student evaluations to initiate changes in instructional practices. However, all changes do not necessarily promote learning (Cohen, 1980). Faculty familiar with SET literature can be expected to make better use of the information than those who are unaware.

Administrative support of programs that help faculty interpret ratings and incorporate different teaching strategies will help ensure the ethics of methods used to increase ratings. Offering consultation with an experienced, sympathetic mentor should encourage pedagogy and morale.

Faculty development is crucial in a time when teaching outcomes and accountability are under scrutiny (Feldman, 1997; Perry & Smart, 1997; Read et al.,

2001; Waters et al., 1988). Both goals, pedagogy and improved SET ratings, should be facilitated through faculty development programs. Faculty often have little or no formal training in teaching (Perry & Smart, 1997). Many do not realize that ratings can be increased and learning increased by simple stylistic changes such as speaking more enthusiastically. Development programs that address pedagogical improvement should strengthen teaching effectiveness and job satisfaction.

The SET process does not seem to affect morale of instructors and faculty respondents report that they use information from evaluations to alter instructional practices. Therefore it should be possible for administrators to use the evaluation process to improve teaching knowing that faculty confidence will be preserved. Instructors should welcome practices and procedures that encourage progressive development of education and experimentation with diverse teaching approaches. Additional development of in-depth questionnaires should provide both information on the efficacy of teaching styles and innovative ideas to be tested. Certainly, faculty input into the development of questions that better address their concerns would provide benefit the advancement of teaching.

On the other hand, since SET information does not affect morale, it should indicate a disquieting condition in higher education. Self-confidence is raised when efforts are recognized. Effective teachers should be encouraged by favorable student evaluations and should address the issues of a poor evaluation proactively. Less diligent teachers would, at best, ignore the indicators and continue weak teaching practices or, at worst, decrease class requirements. If the morale of instructors is untouched by student ratings, grade inflation and other deleterious practices may reduce learning. In a highly

competitive market, bolstering faculty morale should help institutions become centers of academic excellence that are more attractive to prospective students.

In the future, higher education administrators will require objective, quantifiable evidence as they compete for government money (Trout, 1997). With the decline of confidence in and economic support for higher education (Johnsrud, 1996; Read et al., 2001), administrators face the daunting task of insuring that teaching is the primary task of colleges and universities (Boyer, 1990). Studies such as this one reviewing the evaluation process would demonstrate such a commitment to teaching.

Generalizability of Findings

Several factors should be considered when generalizing the results of this study. The sample population is distinctive in higher education because participants were selected from faculty of public colleges in West Virginia that share the mission of teaching. Certainly, dynamics differ in public and private institutions of higher education. Universities' missions encompass teaching, research, and publication and may not support the teaching mission with the same weight as the colleges in the study.

Because many of the colleges in West Virginia are rural, they have unique cultural expectations. Responses to the *Gall FRET* questionnaire should vary with the pace of living, the economy and job availability. Repeated research in targeted areas of the United States would help determine if acceptance of SET information is indeed common among college and university faculty. It would gauge regional differences in faculty acceptance of evaluations by students.

Institutional policies and procedures should influence the opinions of faculty in West Virginia's public colleges. Conceivably, administrators have created an environment that encourages faculty to experiment with different approaches. Faculty may feel willing to consider student opinions seriously to improve job security in a state where job opportunities are limited. The current financial pressures in higher education may have fostered the era of the student as a customer and influence instructors to consider how to make students happy, which may or may not improve learning.

Recommendations for Further Study

Student evaluations of teaching appear to be institutionalized into higher education with 98% of institutions using the process. The massive body of research on the SET process has established validity, reliability, and potential biases of teacher characteristics, course characteristics, student characteristics and environmental characteristics. Sufficient empirical evidence about attitudes and possible consequences to education's purpose is missing, however. With the increasingly competitive environment and its accompanying financial strain, higher education should benefit from further research into the potentially significant consequences of student evaluations on faculty performance and morale.

Further quantitative and qualitative research in two areas would help make evaluations more meaningful. It must be determined that changes in instructional practices are constructive and not a ploy to attain better student ratings. Secondly, clarifying and separating the purposes of student evaluations would make evaluations more valuable for both administrators and faculty members.

Faculty report using SET information when modifying instructional practices. However, with the very small number of studies addressing the consequences of student evaluations, administrators have no way of knowing if the changes encourage learning. Warranted or not, charges of grade inflation contribute to the censure of higher education. Because the purpose of education at risk, further study into the nature of changes to instructional practices is certainly in order. Since faculty reported that morale was not affected by the evaluations process, research can advance without the fear of demoralizing faculty.

Agreement on the intended purpose of student evaluations is key to developing more effective evaluations (Rifkin, 1995). Information should be more beneficial if separate, more in-depth, evaluations were used for formative and summative purposes. Research into valid and reliable questions and methods addressing each objective would bolster both the purpose and mechanisms of higher education.

A third evaluation possibility would serve to guide student selection of institutions and classes. The evaluation would supply information to aid students in choosing the best educational experiences for their personal goals that should promote more successful learning.

Faculty in this study reported that SET information is important to them. Encouraging faculty involvement in the development of questions addressing their concerns about instructional practices would lead to better use of student evaluation information. Perhaps it would be effective to create a bank of questions, tested for validity and reliability, from which instructors could choose.

Impetus to change how teaching effectiveness is measured could come from further research. Some would argue that with current time and fiscal restraints, resources for expanded inquiry are limited. The current emphasis on teaching in higher education supports this use of resources, however. Possibly individual doctoral and faculty research can fulfill the need.

While some contend that higher education would be better off without student evaluations of teaching (Greenwald & Gilmore, 1997), higher education's historical teaching mandate would be advanced by investigating factors related to instructional practices and learning. Since student evaluations of teaching are an integral part of the system, further research into how to make them work to improve education is important.

References

- Aleamoni, L. M. (1987). Typical faculty concerns about student evaluations of teaching. *New Directions for Teaching and Learning*, 31, 25-31.
- Baxter, E. P. (1991). The TEVAL experience, 1983-88: The impact of a student evaluation of teaching scheme on university teachers. *Studies in Higher Education*, 16(2), 151-179.
- Bernardin, H. J., Dahmus, S. A., & Redmon, G. (1993). Attitudes of first-line supervisors towards subordinate appraisals. *Human Resource Management*, 32(2-3), 315-324.
- Blau, G. (1999). Testing the longitudinal impact of work variables and performance appraisal satisfaction on subsequent overall job satisfaction. *Human Relations*, 52(8), 1099-1113.
- Boyer, E. L. (1990). *Scholarship Reconsidered: Priorities of the Professoriate*. The Carnegie Foundation for the Advancement of Teaching. New Jersey: California/Princeton Fulfillment Services.
- Budig, G. A. (1986). Presidential opinions on today's key issues. *AGB Reports*, 28(3), 22-23.
- Calderon, T. G., & Green, B. P. (1997). Use of multiple information types in assessing accounting faculty teaching performance. *Journal of Accounting Education*, 15(2), 221-239.
- Campbell, D. T., & Stanley, J. C. (1963). *Experimental and Quasi-Experimental Designs for Research*. Boston: Houghton Mifflin Company.
- Carnegie Forum on Education and the Economy. (1986). *A Nation prepared: Teachers for the 21st Century*. Washington, D. C.: Carnegie Forum on Education and the Economy.
- Carnegie Foundation for the Advancement of Teaching (1989). National Survey. Washington, D. C.: Carnegie Foundation for the Advancement of Teaching.
- Carnegie Foundation (2003). Retrieved March 1, 2003, from Carnegiefoundation.org/Classification/CIHE2000/background.html
- Carroll, J. (2002). Getting good teaching evaluations without stand-up comedy. *The Chronicle of Higher Education*. Retrieved April 4, 2002, from <http://chronicle.com/jobs/2002/04/2002041501c.htm>
- Cashin, W. E. (1988). Student ratings of teaching: A summary of the research. IDEA Paper No. 20. (ED302567)

- Centra, J. A. (1972). *What instructors learn from students: Student and faculty ratings of instruction*. Paper presented at the American Educational Research Annual Meeting, Chicago, April, 1972. (HE003161)
- Centra, J. A. (1993). *Reflective faculty evaluation*. San Francisco: Jossey-Bass.
- Centra, J. A. (1994). The use of teaching portfolio and student evaluations for summative evaluation. *Journal of Higher Education*, 65(5), 555-570.
- Clancy, S. G. (1997). STJs on change: Resistance, reaction, or misunderstanding? In C. Fitzgerald & L. K. Kirby (Eds.), *Developing leaders: Research and applications in psychological type and leadership development* (pp. 415-438). Palo Alto: CA: Davies-Black.
- D'Apollonia, S., & Abrami, P. C. (1997). Navigating student ratings of instruction. *The American Psychologist*, 52(11), 1198-1208.
- DeCarlo, T. E., & Leigh, T. W. (1996, April). Impact of salesperson attraction on sales managers' attributions and feedback. *Journal of Marketing*, 60, 47-66.
- Des Jarlais, Christine D. (1996). Determinants of faculty morale in an American public research university: Gender differences within and between academic departments. *Dissertation Abstracts International*, 56(10-A), 3847. (UMI No. 9604146)
- Dillman, D. D. (1978). *Mail and Telephone Surveys*. New York, NY: John Wiley & Sons.
- Edwards, C. H. (2000). Grade inflation: The effects on education quality and personal well being. *Education*, 120(3), 538-547.
- Feldman, K. A. (1997). Identifying exemplary teachers and teaching: Evidence from student ratings. *Effective Teaching in Higher Education: Research and Practice*. 241-320). New York: Agathon Press.
- Fitzgerald, C. (1997). Type development and leadership development: Integrating reality and vision, mind and heart. In C. Fitzgerald & L. K. Kirby (Eds.), *Developing leaders: Research and applications in psychological type and leadership development* (pp. 311-335). Palo alto: CA: Davies- Black.
- Foote, E. (1998). Faculty evaluation in community colleges. *Community College Journal of Research & Practices*, 22(7), 693-697.
- Franklin, J., & Theall, M. (1989). *Who reads ratings: Knowledge, attitude and practice of users of student ratings of instruction*. Paper presented at the annual meeting of the American Educational Research Association. San Francisco, CA, March 27-31.

- Frost, P. J., & Fukami, C.V. (1997). Teaching effectiveness in the organizational science: Recognizing and enhancing the scholarship of teaching. *Academy of Management Journal*, 40(6), 1271-1281.
- Gabris, G. T., & Mitchell, K. (1988). The impact of merit raise scores on employee attitudes: The Matthew Effect of performance appraisal. *Public Personnel Management*, 17(4), 369-385.
- Gage, N. L. (1973). *Mandated Evaluation of Educators: A conference on California's Stull Act*. Stanford, CA, October 12-14, 1972.
- Gardner, J. W. (1990). *On Leadership*. New York, NY: The Free press.
- Gomez-Meija, L. R., & Balkin, D. B. (1992). Determinants of faculty pay: An agency theory perspective. *Academy of Management Journal*, 35(5), 921-955.
- Greenwald, A. G. (1998). How useful are student ratings?: Reactions to issues in the current issues section. *American Psychologist*, 5(11), 1228-1229.
- Greenwald, A. G., & Gillmore, G. M. (1997). Grading leniency is a removable contaminant of student ratings. *American Psychologist*, 52(11), 1209-1217.
- Halford, A. (1984). *Faculty morale-Enhancing it in spite of diminishing resources and challenges*. Paper presented to the International Conference for Community College Chairs, Deans, and Other Instructional Leaders, Phoenix, AZ. (ERIC Document Reproductive Service No. ED 368422)
- Hobson, S. M., & Talbot, D. M. (2001). Understanding student evaluations. *College Teaching*, 49(1), 26-32.
- Johnson, B., & Christensen, L. (2000). *Educational Research*. Needham Heights, MA: Allyn & Bacon.
- Johnsson, C., & Kelley, M. (1998). Productivity and competence: Post-tenure review in American universities. *Issues in Accounting Education*, 13(3), 747-760.
- Johnsrud, L. K. (1996). *Maintaining Morale: A guide to assessing the morale of midlevel administrators and faculty*. The College and University Personnel Association: Washington, DC.
- Kang, B., & Miller, T. (2000). Faculty development: Research findings, the literature base, and directions for future scholarship. *Information Analysis* 070.
- Kerlinger, F. N., & Lee, H. B. (2000). *Foundations of Behavioral Research*, 4th ed. Philadelphia: Harcourt College Publishers.

- Kohli, A. K., & Joworski, B. J. (1994). The influence of coworker feedback on salespeople. *Journal of Marketing*, 58, 82-94.
- Kouzes, J. M., & Pozner, B. Z. (1987). *The Leadership Challenge*. San Francisco, CA: Jossey-Bass.
- Krautman, A. C., & Sanders, W. (1999). Grades and student evaluations of teachers. *Economics of Education Review*, 18, 59-63.
- Kulik, J. A., & McKeachie, W. J. (1976). The evaluation of teachers in higher education. In F. N. Kerlinger (Ed.). *Review of Research in Higher Education*, 3. Itasca, Ill.: Peacock.
- Lanning, W., & Perkins, P. (1995). Grade inflation: A consideration of additional causes. *Journal of Instructional Psychology*, 22(2), 163-169.
- Manger, D. K. (1997). Report says standards used to evaluate research should also be used for teaching and service. *Chronicle of Higher Education*, 44(2), A18-A19.
- Marlin, J. W. (1987). Student perceptions of end of course evaluations. *Journal of Higher Education*, 58, 704-716.
- Marsh, H. W. (1984). Students' evaluations of university teaching: Dimensionality, reliability, validity, potential biases, and utility. *Journal of Educational Psychology*, 76(5), 707-754.
- Marsh, H. W. (1987). Student's evaluation of university teaching; Research findings, methodological issues, and directions for further research. *International Journal of Educational Research*, 11, 253-388.
- Marsh, H. W., & Dunkin, M. J. (1992). Students' evaluation of university teaching: A multidimensional perspective. In R. P. Perry & J. C. Smart (Eds.), *Effective Teaching in Higher Education: Research and Practice* (pp. 241-320). New York: Agathon Press.
- Marsh, H. W., & Roche, L. A. (1997). Making students' evaluations of teaching effectiveness effective. *American Psychologist*, 52(11), 1187-1197.
- Martin, J. R. (1998, November). Evaluating faculty based on student opinions: Problems, implications, and recommendations from Deming's theory of management perspective. *Issues in Accounting Education*, 1079-1094.
- McCullough, M. A., & Gremler, D. D. (1999). Guaranteeing student satisfaction: An exercise in treating students as customers. *Journal of Marketing Education*, 21(2), 118-132.

- McKeachie, W. J. (1979). Student rating of faculty: A reprise. *Academe*, 65, 384-397.
- McKeachie, W. J. (1990). Research on college teaching: The historical background. *Journal of Educational Psychology*, 82(2), 189-200.
- McKeachie, W. J. (1997). Student ratings. *American Psychologist*, 52(11), 1218-1226.
- McKenzie, R. B. (1975). The economic effects of grade inflation on instructor evaluations: A theoretical approach. *Journal of Economic Education*, 6(2), 99-105.
- Millman, J. (1981). *Handbook of Teacher Evaluation*. Sage Publications: Beverly Hills, London.
- Odden, Allan (2000). *Motivational Theory*. Retrieved November 4, 2002, from University of Wisconsin-Madison, Consortium for Policy Research in Education Web site: www.wcer.wisc.edu/cpre/tcomp/research/strategy/motivational.asp
- Pearce, J. T., & Porter, L. W. (1986). Employee responses to formal performance appraisal feedback. *Journal of Applied Psychology*, 71, 211-218.
- Perry, R. P., & Smart, J. C. (Eds.). (1997). *Effective Teaching in Higher Education*. Agathon Press: New York.
- Radmacher, S. A., & Martin, D. J. (2001). Identifying significant predictions of student evaluations of faculty through hierarchical regression analysis. *Journal of Psychology*, 135(3), 259-269.
- 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), 189-193.
- Remmers, H. H. (1928). The relationship between students' marks and students' attitudes towards instructors. *School and Society*, 28, 759-760.
- Rifkin, T. (1995). *The status and scope of faculty evaluation*. Washington, DC: Office of Educational Research and Improvement (ERIC Digest No. ED385315).
- Rosen, R. H., & Brown, P. B. (1996). *Leading People*. New York, NY: Viking.
- Rutland, P. (1990). Some considerations regarding teaching evaluations. *Political Science Teacher*, 3(4), 1-2.
- Ryan, J. J., Anderson, J. A., & Birchler, A. B. (1980). Student evaluations: The faculty responds. *Research in Higher Education*, 12(4), 317-333.

- Seldin, P. (1980). *Successful Faculty Evaluation Programs: A Practical Guide to Improve Faculty Performance and Promotion /Tenure Decisions*. Coventry Press: Crugers, NY.
- Seldin, P. (1984). *Changing practices in faculty evaluation*. San Francisco: Jossey-Bass.
- Seldin, P. (1993). The use and abuse of student ratings of professors. *Chronicle of Higher Education*, 39(46), A40.
- Simpson, P. M., & Siguaw, J. A. (2000). Student evaluations of teaching: An exploratory study of the faculty response. *Journal of Marketing Education*, 22(3), 199-214.
- Spencer, P. A., & Flyr., M. L. (1992). *The formal evaluation as an impetus to classroom change: Myth or reality*. Unpublished master's thesis. Riverside Community College, University of California, Riverside, United States.
- Stapleton, R. J., & Murkison, G. (2001). Optimizing the fairness of student evaluations: A study of correlations between instructor excellence, study production, learning production, and expected grades. *Journal of Management Education*, 25(3), 269-291.
- Stratton, R. W., and Myers, S. C. (1994). Faculty behavior, grades, and student Evaluations. *Journal of Economic Education*, 25(1), 5-16.
- Swenson, C. (1998). Customers and markets: The cuss words of academe. *Change*, 30(5), 34-39.
- Top Education (2001). *Theories of Motivation*. Retrieved November 4, 2002, from [http://www.top-education.com/Management/motivation theories 2 asp](http://www.top-education.com/Management/motivation%20theories%20asp)
- Trout, P. A. (1997). What the numbers mean. *Change*, 29(5), 24-30.
- Trout, P. (2000). Teacher evaluations. *Commonweal*, 138(8). 10-12.
- Vroom, V. H. (1964). *Work and Motivation*. John Wiley & Sons, Inc: New York, London, Sydney.
- Wachtel, H. K. (1998). Student Evaluations of College Teaching Effectiveness: A Brief Review. *Assessments & Evaluation in Higher Education*, 23(2), 191-212.
- Wallace, J. J., & Wallace, W. A. (1998). Why the costs of student evaluations have long since exceeded their value. *Issues in Accounting Education*, 13(2), 443-448.
- Waters, M., Kemp, E., & Pucci, A. (1988). High and low faculty evaluations: Descriptions by students. *Teaching of Psychology*, 15(4), 203.

- Williams, W. M., & Ceci, S. J. (1997). 'How'm I doing?': Problems with student ratings of instructors and courses. *Change*, 29(5).
- Wilson, R. (1998). New research casts doubt on value of student evaluations of professors. *Chronicle of Higher Education*, 44(19), A12-A14.
- Zelby, L. W. (1974). Student-faculty evaluation. *Science*, 183(4131), 1267-1270.
- Zoller, U. (1992). Faculty teaching performance evaluations in higher science education: Issues and implications. *Science Education*, 76(6), 673-684.

Appendix A

Faculty Response to Evaluations of Teaching

Faculty Response to Evaluations of Teaching

This survey is designed to elicit individual faculty reaction to student evaluations of teaching used at public colleges in West Virginia. If you wish to comment on any question or qualify your answers, please feel free to use the back of the questionnaire. Your comments will be taken into account

Please fill in or checkmark the box indicating your considered response for every question.

1. **Sex:** Female Male
2. **Faculty Status:** Tenure Track- Tenured Tenure Track- Not-Tenured Not Tenure Track Adjunct/Part-Time
3. **College/School:** Applied Sciences Education Business Humanities Social Sciences Fine Arts Other
4. **Ethnicity:** Am. Indian/ Alaska Native Asian Native Hawaiian/ Other Pacific Islander
 Black/African American Hispanic/Latino White Other _____
5. **Years of Higher Education Teaching Experience:** _____ yrs
6. **Age:** _____ yrs
7. **Years at Present Institution:** _____ yrs

Please checkmark your answer to the following questions on a continuum:
 Strongly Agree, Agree, Disagree or Strongly Disagree

The following questions concern your opinions of student evaluations.	Strongly Agree	Agree	Disagree	Strongly Disagree
8 Students are capable of evaluating teaching.	<input checked="" type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>
9 Mandated student evaluations of teaching provide students with an appropriate vehicle to evaluate teaching.	<input checked="" type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>
1 Student evaluations of teaching should be used by instructors to improve teaching.	<input checked="" type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>
1 Student evaluations of teaching should be used for administrative decisions related to retention, promotion, and tenure.	<input checked="" type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>
1 Professors can raise student ratings of teaching by making a concentrated effort to develop more effective instructional practices.	<input checked="" type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>
1 Professors can raise evaluation ratings through instructional activities that please students without necessarily enhancing student achievement.	<input checked="" type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>
1 Generally speaking, student evaluations of teaching affect faculty morale.	<input checked="" type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>
1 Student evaluations of teaching affect instructors' self-images as educators.	<input checked="" type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>
1 Student evaluations of teaching affect the satisfaction instructors derive from teaching.	<input checked="" type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>

Faculty Response to Evaluations of Teaching

This survey is designed to elicit individual faculty reaction to student evaluations of teaching used at public colleges in West Virginia. If you wish to comment on any question or qualify your answers, please feel free to use the back of the questionnaire. Your comments will be taken into account

The information I have gleaned from student evaluations of my teaching has encouraged me to validate or change ...		Strongly Agree	Agree	Disagree	Strongly Disagree
1 7course objectives.	9	9	9	9
1 8the use of group discussion in my classes.	9	9	9	9
1 9the amount of lecturing I use.	9	9	9	9
2 0the amount of handouts or other course aids.	9	9	9	9
2 1the process of returning exams and papers.	9	9	9	9
2 2the content of my classes.	9	9	9	9
2 3the amount of material covered in a class.	9	9	9	9
2 4the relevance of class material to student interests.	9	9	9	9
2 5the difficulty level of my classes.	9	9	9	9
2 6the way I respond to students' questions.	9	9	9	9
2 7the way I interact with students outside of class.	9	9	9	9
2 8my grading standards.	9	9	9	9
2 9	I use student feedback to evaluate my teaching.	9	9	9	9
3 0	I examine results of institution-wide student evaluations of teaching	9	9	9	9

Your contribution to this effort is very greatly appreciated. If you would like a summary of results, please print your name and address on the back of the return envelope (not questionnaire) or send email request to annettegall@charter.net We will see that it gets to you. Our return address is: Marshall University Leadership Studies, PO Box 2547, Chas., WV 25329.

Appendix B
First Cover Letter

Marshall University Graduate College
College of Human Resources and Education
Educational Leadership Department
100 Angus E. Peyton Drive
South Charleston, WV 25303

Dear Faculty Member,

Marshall University is sponsoring a doctoral study on the perceived effects of Student Evaluations of Teaching (SET) on instructional practices and instructor morale. Despite the volumes of research literature and nearly universal use of the SET process in higher education, very few studies investigate how you, the faculty member, regard the process and how the evaluations might affect your primary task of teaching. This is an opportunity for the faculty to voice opinions of the evaluation process to either support strengthening the process or changing it to better serve your needs.

The enclosed one page questionnaire will be sent to a randomly selected sample of the faculties of the eight, public colleges in West Virginia. A reminder note and second mailing of the questionnaire may follow to encourage a larger return rate. We understand that you have many pressing issues to address in the course of your day, but we are confident that the results will be important to our profession. The questionnaire was designed to take no more than ten minutes to complete.

Participation is entirely voluntary and your specific contributions will be anonymous. To ensure complete confidentiality, please do not write your name anywhere on the questionnaire. Results will be available by request.

We greatly appreciate your time and help with this project. If you have any questions or want a copy of the results, please e-mail annettegall@charter.net or call 1-304-343-9697.

Very truly yours,

Annette R. Gall , M.S., Ed.S.

Appendix C
Reminder Postcard

Date

Last week a questionnaire seeking your opinion about how student evaluations of teaching effect instruction and morale was mailed to you. Your name was drawn in a random sample of faculty from West Virginia's public colleges.

If you have already completed and returned it to us please accept our sincere thanks. If not, please do so today. Because it has been sent only to a small, but representative sample of faculty, it is extremely important that yours also be included in the study is the results are to accurately represent the opinions of faculty in West Virginia.

If by some chance you did not receive the questionnaire, or it got misplaced, please email me right now and I will get another one in the mail to you today.

Sincerely,

Annette R. Gall
Project Director

Appendix D

Follow-up Cover Letter

September 29, 2003

Dear Faculty Member,

About three weeks ago, I wrote to you seeking your opinion on the effects of student evaluations of teaching on instructional practices and faculty morale. If you have returned the questionnaire, please accept our sincere thanks.

The opinions of faculty members are crucial in deciding whether student evaluations of teaching decrease or increase morale and instructional practices. We undertook this research because it is our belief that faculty are the heart of higher education and anything that weakens your commitment undermines the primary task of teaching.

I write to you again because of the significance each questionnaire has to the usefulness of this study. You were chosen through a scientific sampling process in which every faculty member of public colleges in West Virginia had an equal chance of being selected. This means that only one out of every 3.4 faculty members is being asked to complete this questionnaire. In order for the results of this study to be truly representative of the opinions of the faculties, it is essential that each person in the sample return the questionnaire.

A replacement questionnaire and self addressed stamped envelope is enclosed. Please complete and return it as quickly as possible.

I'll be happy to send you a copy of the results should you so desire. Simply put your name, address, and "copy of results requested" on the back on the return envelope. This will not compromise your anonymity. We expect to have the completed document by early next spring.

Your contribution to the success of this study will be appreciated greatly.

Most sincerely,

Annette R. Gall, M.S., Ed.S.