

Lynn University

SPIRAL

Student Theses, Dissertations, Portfolios and
Projects

Theses and Dissertations Collections

4-29-2008

A Study of Faculty Perceptions of the Use of Student Evaluations of Teaching in Faculty Assessments, Promotion and Tenure Decisions

Judith Kanegita Cineas
Lynn University

Follow this and additional works at: <https://spiral.lynn.edu/etds>

Recommended Citation

Cineas, Judith Kanegita, "A Study of Faculty Perceptions of the Use of Student Evaluations of Teaching in Faculty Assessments, Promotion and Tenure Decisions" (2008). *Student Theses, Dissertations, Portfolios and Projects*. 78.

<https://spiral.lynn.edu/etds/78>

This Dissertation is brought to you for free and open access by the Theses and Dissertations Collections at SPIRAL. It has been accepted for inclusion in Student Theses, Dissertations, Portfolios and Projects by an authorized administrator of SPIRAL. For more information, please contact liadarola@lynn.edu.

**A Study of Faculty Perceptions of the Use of Student Evaluations of Teaching in
Faculty Assessments, Promotion and Tenure Decisions**

DISSERTATION

**Presented in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy
Lynn University**

**By
Judith Kanégita Cinéas**

**Lynn University
April 29, 2008**

**LYNN UNIVERSITY
LIBRARY**

ACKNOWLEDGEMENTS

WOW... I am truly blessed to have been surrounded by the many people who have each in their own way contributed to my ability to attain this milestone.

Dr. Crawford and Dr. Levenson, I could not have chosen a better committee if I tried. Thank you for all the support and guidance. Dr. Cipolla, there are no words. You have been more than amazing. Without your guidance I could not have completed this dissertation.

Many thanks to my friends who have been constant supporters and motivators. Thank you for your confidence.

My dear friends who have been with me through the dissertation process. Dr. Régine Joseph, I remember us talking about our doctorates in Middle School. Ms. Eunide Vaval, your support has been invaluable. Going through this process with friends has certainly enhanced the experience. This is our year!

Without the love and support of my family, this experience would have been much more difficult. My wonderful brothers and sisters, who are always proud of me and continue to inspire me, especially Didier Elysee and Anick Elysee. Thank you for putting up with all that you have, particularly in these recent months.

In all my years, through all my studies, I still have not found the right words and certainly not enough them, to thank my mother, Marie Nicole Cinéas. It has been such a blessing to have the love and constant encouragement of my family. The greatest blessing of all is my mother.

Abstract

The relationship between student evaluations of teaching (SET), grade inflation, and faculty assessments, promotion and tenure (FAPT) decisions is an issue that has received much attention in the past 30-40 years. The literature speculates that the use of SET ratings in FAPT decisions has resulted in faculty giving students higher grades to secure higher SET. This trend, it is held, has contributed to grade inflation, which decreases the reliability of assessing teaching effectiveness using student grades since they are not reflective of actual learning. It also decreases the validity of SET as an evaluation method if inflated grades are yielding higher SET ratings. This study explored the faculty perception of these uses of SET and its impact on grade distribution. The purpose of this study was to determine faculty perceptions regarding the uses of SET in FAPT decisions. The results of this study provide empirical information for administrators in higher education to evaluate the use of SET as an assessment method.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	ii
ABSTRACT	iii
TABLE OF CONTENTS	iv
LIST OF TABLES	vi
LIST OF FIGURES	vii
CHAPTER ONE: INTRODUCTION TO THE STUDY	1
<i>Purpose of the Study</i>	3
<i>Definition of Terms</i>	3
<i>Scope</i>	4
<i>Interest Significance and rationale</i>	7
<i>Research Questions and Hypotheses</i>	9
CHAPTER TWO: REVIEW OF THE LITERATURE	12
Student Evaluations of Teaching (SET)	
<i>Purpose of SET</i>	12
<i>Factors influencing SET</i>	13
<i>Reliability and Validity of SET</i>	21
<i>Student Evaluations of Faculty Performance Versus Student evaluations of Teaching Strategies</i>	22
Grade inflation	23
<i>Definition</i>	23
<i>Theories about the causes</i>	24
<i>Impact (negative) of Grade Inflation on Students and Institutional Accountability</i>	29
<i>Grade Inflation and Student Evaluations of Teaching (Studies)</i>	30
<i>Trends, Controversy, and Emerging Strategies to Manage Grade Inflation</i>	32
<i>Faculty Assessments, Promotion and Tenure, (FAPT) Decisions</i>	35
Discussion of the Literature	
<i>Theoretical Literature</i>	37
<i>Empirical Literature</i>	38
<i>Conclusion</i>	39
CHAPTER THREE: RESEARCH DESIGN AND PROCEDURE	
<i>Research Methodology</i>	41
<i>Research Design</i>	41
<i>Population and Sample</i>	42
<i>Evaluation of Sampling Design</i>	43
<i>Data Collection Procedure</i>	45
<i>Instruments</i>	45

<i>Data Analysis</i>	46
CHAPTER: RESULTS	51
Research Question 1: <i>What are the faculty perceptions of uses of Student Evaluations of Teaching?</i>	55
Research Question 2: <i>How much impact do faculty believe that Student Evaluations of Teaching have on FAPT decisions?</i>	59
Research Question 3: <i>Do faculty believe that Student Evaluations of Teaching have an impact on Grade Distribution?</i>	61
Research Question 4: <i>Are faculty interested in a unified grading system? If so, who should be responsible for it??</i>	64
Hypothesis 1: <i>ICUF faculty will experience greater pressure to have high SET ratings than SUS faculty</i>	65
Hypothesis 2: <i>SUS faculty will report less impact of SET on their grading practices than their ICUF counterparts</i>	68
Hypothesis 3: <i>Class size will be inversely related to perceived importance of SET ratings</i>	70
Hypothesis 4: <i>Faculty who perceive the importance of SET on FAPT to be high are more likely to consider SET ratings in preparing student assignments and grading</i>	72
CHAPTER FIVE: DISCUSSION OF FINDINGS	
<i>Interpretations</i>	76
<i>Practical Implications</i>	84
<i>Conclusions</i>	86
<i>Strengths of the study</i>	88
<i>Limitations</i>	89
<i>Recommendations for Future Studies</i>	90
REFERENCES	94
BIBLIOGRAPHY	103
APPENDICES	
Appendix A: <i>IRB Approval Letter</i>	105
Appendix B: <i>Introductory Email to Office of Academic Affairs/Provost</i>	107
Appendix C: <i>Email Request to Office of Academic Affairs/Provosts</i>	109
Appendix D: <i>Invitation to Participate Sent to Professors</i>	111
Appendix E: <i>Follow up Email Sent to Professors</i>	113
Appendix F: <i>Copy of Electronic Survey (including Consent form)</i>	115
VITA	129

LIST OF TABLES

	Page
Table 1: <i>Demographic Characteristics</i>	54
Table 2: <i>Uses of Student Evaluations of Teaching</i>	57
Table 3: <i>Perceptions of Student Evaluations of Teaching</i>	58
Table 4: <i>Awareness of uses of SET</i>	60
Table 5: <i>Pearson Correlation between variables</i>	63
Table 6: <i>Perceived SET Impact on Grading</i>	67
Table 7: <i>Perceived Weight of SET Rating</i>	69
Table 8: <i>Class size correlations</i>	71
Table 9: <i>Pearson Correlation between perceived importance of SET and faculty behaviors</i>	73

Introduction

In the last 40 years, there has been a noticeable change in the grades received by students enrolled in institutions of higher education. This trend includes more students receiving grades of A- or higher, and fewer students receiving grades of C or lower (Nagle, 1998; McSpirit, Kopacz, Jones, & Chapman, 2000). Universities have been faced with the issue of grade inflation for the past 40 years (Nagle, 1998). Initially, one of the main reasons for this change was to help students avoid being drafted at the time of the Vietnam War (Birnbaum, 1977; Voge & Higbee, 2004). In today's colleges and universities, instructors continue to face a great deal of pressure to award higher grades. The continued pressure is in part, because the grades students receive or expect to receive can have great impact on whether students enroll in future classes taught by the instructor (Voge & Higbee, 2004), as well as their ratings on student evaluations of teaching (SET) (Baldwin, 2004; Ellis, Burke, Lomire, & McCormack, 2003; Greenwald & Gilmore, 1997; Grimes et al., 2004; Isely & Singh, 2005; Millea & Grimes, 2002). Researchers have noted an increased use of student evaluations in the faculty assessments, promotion and tenure (FAPT) decisions. The researchers assert that instructors have responded to this change by inflating student grades (Algozinne, Beattie, Bray, Flowers, Gretes, Howley et al. 2004; Yunker & Marlin, 1984).

Voge and Higbee (2004) quote a Duke University professor who said, "A's are as common as dirt in universities nowadays" (p. 64) as he explained that a higher percentage of students now receive A's. Voge and Higbee (2004) report that a professor admitted that if instructors were giving students the grades they deserve, enrollment in their classes would decrease in the years to come. Another professor at Harvard University described

this as a situation where professors assign grades based not on student performance but on the desired faculty evaluation ratings (Voge & Higbee, 2004). As a result, there is a continuous debate on whether students in colleges and universities are earning higher grades than students did ten or twenty years ago, or if they are simply receiving higher grades which are not based upon ability or performance (Hanson, 1998).

Researchers (e.g., Birnbaum, 2000) have suggested that students may not be prepared (McAlpine & Harris, 2002), or may not possess the expertise necessary to effectively evaluate the quality of teaching. The added responsibility of evaluating instructors can therefore place students in a situation where their need for favorable grades is being met at the expense of the knowledge they should be acquiring. This bargain is based upon the exchange of higher grades for higher ratings (Birnbaum, 1977; Scanlan & Care, 2004). Grade inflation and the education of college and university students is a concern for students, educators, and certainly for the organizations which will employ them, as well as the individuals who will receive services from them in the future. Institutions of higher education throughout the nation face concerns regarding grade inflation. This study is relevant in that it explored the issue from the faculty's perspective. Moreover, it sought to explore the possible impact that one method of evaluation can have on those being evaluated, as well as the evaluators.

The use of student evaluations of teaching as a method of assessing classroom teaching is becoming a problem, as it may be compromising the quality of higher education. When faculty resign to giving higher grades to get higher SET ratings, the students may not be challenged to perform at an academic level becoming of college students. Students then may graduate with high marks, but still be unable to perform in

the workforce at the level indicated by their academic history since they did not acquire that level of knowledge. Since students are rating professors based on the grades they anticipate receiving in the class, increases in the pressure to have higher ratings on SET are likely to result in a decrease in rigor of demands placed on college students (Birnbaum, 2000). Lowered grading standards and performance requirements are likely to result in the graduation of unprepared students becoming unprepared service providers.

Purpose of the Study

The purpose of this study was to explore and determine whether there is a significant association between the use of student evaluations of teaching in FAPT decisions and faculty behaviors, including grade inflation. The study explored faculty's perceptions of students' evaluation of classroom teaching, and faculty's beliefs about the impact these documents have on FAPT decisions. The study also investigated the impact that faculty's perception of the use of these documents in FAPT decisions has on the instructors' behaviors, including grading practices. Faculty's agreement with the current use of student evaluations of teaching, as well as recommendations for resolving issues with grade inflation were examined.

Definition of terms

Grade inflation refers to "a rise in grades that is unwarranted by student performance" (Voge & Higbee, 2004. p. 65). *Grade-rationing* policies limit the percentage of students who may get A's in a course (Roarty, 2004). *Student evaluations of teaching* (SET) provide students the opportunity to assess the instructor and classroom teaching. While this is, in essence, a grade for the instructor, it is merely the opinion of students who rate "satisfaction" with teaching by instructors (Algozinne et al., 2004).

Teaching performance is a method of calculating the success of a teaching technique (AAUP, 2001). *Teaching effectiveness* refers to the instructor's ability to manage the time to present material and deliver instruction in a manner that facilitates the students' ability to learn (Yunker & Marlin, 1984). This also involves creating an environment that increases the students' knowledge or understanding (Buck, 1998). *Student learning*, on the other hand, measures the efficacy of the teaching methodology (AAUP, 2001) and is based on outcomes, performance, or achievement. *Faculty Assessment, promotion and tenure decisions* (FAPT) generally result from an annual or periodic review of faculty activities in the areas of teaching, service, scholarship and/or research. Student evaluations of teaching plays an important part in this review (Algozinne et al., 2004; Yunker & Marlin, 1984). The *Independent Colleges and Universities of Florida (ICUF)* is an association consisting of 28 private institutions. The *State University System (SUS)* is comprised of Florida's eleven state universities.

Scope

The study explored the relationship between the administrative uses of SET in FAPT decisions and how this process impacts faculty behaviors in terms of grading. The primary causal variable is student evaluations of teaching. Faculty behaviors in response to the causal variable such as work load reduction and grade inflation are some of the outcome variables that are examined to determine how they are impacted by student evaluations of teaching. Intervening, contextual or mediating variables that further influence the relationship between SET and outcome variables, include other student culture, pedagogical culture and institutional culture, variables such as institutional

grading policies, criteria for reappointment, promotion, tenure, multi-year contracts and other merit decisions, and measures of teaching effectiveness.

The topic for this study of faculty perceptions of toward the use of student evaluations of teaching in faculty assessments, promotion and tenure decisions, was identified after a review of the literature revealed the increasing debate surrounding the issue. As the concerns about grade inflation become more prominent, universities around the country seek to find ways to remedy the problem. While the concern is focused on higher education, this is an issue which can affect any industry as it relates to the impact an evaluation method [SET] may be having on those being evaluated. More specifically, the concern is the impact of using a measurement tool which lacks validity. The global impact that is presented by this change in behavior [grade inflation] resulting from the use of an evaluation method, makes this a topic of interest not just in education, but in administration at large.

Some questions answered through this study are:

1. Do faculty view SET as a valid measure of teaching effectiveness?
2. To what degree do faculty believe SET are used as a measure in faculty assessments, promotion and tenure decisions?
3. Is there a significant association between the administrative uses of SET and faculty behaviors?

Though the concerns about the relationship between SET and grade inflation are persistent in colleges and universities (Falkenberg, 1996; FAU, 2004; Harvey, 1999; Roarty, 2004; University of Southern Indiana, n.d.; Voge & Higbee, 2004), this is a problem of policies and practices. It relates to the accountability of students and

institutions, the competitive economic issues to retain students, the consumer-student focus of institutions, and the need to explore student, pedagogical and institutional cultures. The concerns with the uses of SET and their possible impact stem from the opportunity and responsibility that students have been extended, to evaluate the effectiveness of teaching methods. These evaluations, which are often used in making decisions regarding faculty salaries, promotion, tenure and reappointment decisions (Parayitam, Desai & Phelps, 2007) may not be reliable or even valid measures of the effectiveness of a professor's teaching. The investigation in this study sought to explore the relationships between SET and grade inflation, with respect to faculty's perception of this use of SET in faculty reappointment, promotion and tenure decisions, and to identify areas of future scholarly inquiry.

The primary concern is the need for an effective method of evaluating the teaching of professors without compromising the integrity of higher education. As it stands, students in higher education are being asked to evaluate the performance of their instructors, the quality of instruction they receive, and instructors' effectiveness in facilitating student learning (Birnbaum, 2000; McAlpine & Harris, 2002). This method of evaluating teaching may result in students having a significant impact on institutional decisions to change the status of faculty. This also potentially places the faculty in the compromising position of having to choose between assigning the grades that students deserve, or preserving their positions by inflating students' grades.

By assigning higher grades to students, professors can increase the likelihood that they will get favorable evaluations (Birnbaum, 1977; Scanlan & Care, 2004). In addition, the types of grades they assign also impact enrollment in courses they teach in the future

(Voge & Higbee, 2004). With increased competition on campuses, professors whose students receive lower grades face the potential of a decline in enrollment in future classes, which for non-tenured faculty is a potential factor in their institution's merit decisions, such as whether to renew or discontinue a professor's contract (Voge & Higbee, 2004).

Context

Higher education is the primary discipline of focus in this study, with an emphasis on job performance evaluations and employees' response behaviors. A more specific area of concentration, is the heavy inclusion SET in faculty assessments. Sociology is another discipline examined, as it applies to the learning environment, and the shift of power to students, that can result from the increased value given to the SET. The literature reviewed spans the years 1972 to 2008. The time span for the review includes earlier literature because there are early institutional policies and research that apply to the current issue of academic evaluation and grade inflation. Literature from the previous ten years is reviewed primarily, with most emphasis on publications from 2000 to present. The literature reviewed was limited to publications in the United States.

Interest, Significance, and Rationale

The education of students in American colleges and universities is a topic that should be of interest to the general population, since the quality of education that students receive can impact the level at which they will be able to become productive members of society. For many students, this is the time when they should be gaining knowledge that will help them flourish in their careers, yet the added responsibility of evaluating instructors can place them in a situation where their desire for favorable grades is being

met while their need to acquire valuable knowledge is compromised. The concern surrounding grade inflation and the education of college and university students is a concern for students, educators, and certainly for the organizations which will employ them in the future. Institutions of higher education throughout the nation face this difficulty. The potential impact of SET as a method of performance evaluation should be troublesome to employers, as it can show the influence that the evaluation itself can have on performance.

While the regional accrediting bodies establish some standards for teaching at the post secondary level and periodically review university teaching plans such as syllabi, individual colleges and universities have always been in charge of establishing their own grading procedures (Western Association of Schools and Colleges [WASC], 2001). With the current practice, at this time universities and colleges are taking their own approaches to resolving the problem of grade inflation. Examination of grade inflation in light of the impact of SET on FAPT decisions needs exploration. A review of the literature is provided in the next section. Theoretical, empirical, and methodological literature is presented in the critical analysis in the next section. The review of literature also includes conclusions and recommendations for future inquiry into faculty perceptions of toward the use of student evaluations of teaching in faculty assessments, promotion and tenure decisions.

In much of the available literature, researchers purport that grade inflation does exist in part as a result of the SET being used for FAPT decisions, leading to somewhat of a quid pro quo between students and professors. There is limited information available that focuses on the professors' view point. Many professors deny that this exchange is

actually occurring. One goal of this research was to explore that avenue by allowing professors to anonymously-self report on their thoughts on the use of SET in FAPT decisions and the impact reported in research to have resulted, and when applicable, their own behaviors or involvements in inflating student grades.

Research Questions

This study attempted to answer the following research questions:

1. What are the faculty perceptions of the uses of Student Evaluations of Teaching?
2. How much impact do faculty believe that Student Evaluations of Teaching have on Faculty Assessments, Promotion and Tenure decisions?
3. Do faculty believe that Student Evaluations of Teaching have an impact on Grade Distribution?
4. Are faculty interested in a unified grading system?

If so, who should be responsible for it?

Hypothesis

1. ICUF faculty will experience greater pressure to have high SET ratings than SUS faculty.
2. SUS faculty will report less impact of SET on their grading practices than their ICUF counterparts.
3. Class size will be inversely related to perceived importance of SET ratings.
4. Faculty who perceive the importance of SET in FAPT to be high are more likely to consider SET ratings in preparing student assignments and grading.

Assumptions

The literature suggests that faculty are interested in promoting increased learning. When faculty believe that current or lowered standards allow them to get the desired SET ratings, they are not able to implement the increased workload and more stringent grading practices which would motivate the increase in student efforts leading to increased learning (Birnbaum, 2000). The assumption then can be made that if faculty do report that they are lowering standards to obtain higher SET ratings, institutions of higher learning will be able to grasp from such admissions that SET are not sufficient to effectively measure teaching effectiveness for the purpose of FAPT decision. A reduction in the emphasis on SET in the FAPT decision could potentially reintroduce the original high standards and quality of learning that were once associated with higher education.

Limitations

The population for this study was comprised of faculty teaching at one of the eleven institutions in Florida's State University System (SUS) or 28 Independent Colleges and Universities of Florida (ICUF). All of these institutions have a Level II or higher accreditation from the Southern Association of Colleges and Schools. Email invitations were sent to the faculty, inviting them to complete the online survey.

Data was only collected in Florida, from the faculty at the eleven SUS or 28 ICUF institutions which have been selected for this study. Therefore, the study would need to be replicated in other regions of the country to increase generalizability. The data collected was entered into Statistical Package for the Social Sciences (SPSS) for analysis.

Another limitation of the study is that the data consisted of information that is self reported by the participants, making it possible for self-report bias to affect data. The participants were also a self-selected group who chose to complete the survey.

CHAPTER II

Literature Review

Purposes of Student Evaluations of Teaching

Centra (1972) and Murray (1997) identified the two most important reasons for the evaluations of teaching as (1) improvement of teaching and (2) an aid in making faculty promotion decisions. Colleges and universities, as well as accrediting bodies, have acknowledged that there is a need for continuous evaluation of the type and quality of instruction that students receive in the classroom (American Association of University Professors [AAUP], 2001).

The University of Southern Indiana (2002) created a task force to address concerns pertaining to SET. The task force was charged with reviewing the current method of evaluation and how it applies to methods of teaching used at the university. The task force determined that SET can be divided in two categories: summative and formative. The task force clarified the purpose of each of the two styles of evaluation. The summative evaluations, often multiple choice style of evaluation, were used for administrative purposes. The formative evaluations, which involved more detail, including open-ended questions for students to answer, were used by the faculty for the purpose of improving the course. In their position statement, the task force recommended that the university adopt an evaluation system that includes both evaluation forms (University of Southern Indiana, 2002).

Student evaluations of teaching are generally administered in the form of a questionnaire that students complete at the end of the course (summative evaluation).

Generally, in this multiple choice questionnaire format of SET, students select from provided responses. Students may be asked to write out evaluations so that the process can include both summative and formative responses (Haskell, 1997). Nummedal (1994) suggested the use of formative evaluation of classroom instruction, a process in which faculty seek student feedback on instruction, while the class is in progress. Utilizing the feedback, faculty can begin to implement observable changes noticeable to the students who can, in turn, assist in determining if the changes are indeed improving the course. This form of evaluation, Nummedal (1994) stated, elicits the students' participation in the instruction process as it forces them to reflect on their own learning and evaluate what they are or are not getting from the classroom instruction.

While many professors have used and continue to use some version of SET as a tool to help improve teaching, this method of evaluating classroom teaching was initiated in the 1920s (Algozinne et al., 2004). Though widely used through the 1960s, this evaluation was primarily voluntary, whereas today it is mandatory in most institutions of higher education. Initially SET were meant to be used as a tool to assist faculty in evaluating the strong and weak points of their instruction, now the SET are being used for a different purpose (Murray, 1997). Administrators in colleges and universities are using the SET in the review of faculty performance, when making human resources decisions. While the initial use was welcomed, this recent use has incited much controversy (Algozinne et al., 2004; Eiszler, 2002; Yunker & Marlin, 1984).

Factors influencing SET

Berg and Lindseth (2004) conducted a study of factors that were affecting students' perception of what constitutes effective teachings versus ineffective teaching

using a convenience sample of 252 participants. The sample included sophomores, juniors and seniors, and excluded freshmen. Participants responded to a three-part questionnaire, intended to assess demographic information, characteristics used to differentiate effective from ineffective teaching and questions pertaining to coursework and grading. The sample also consisted of students from only one university. The study had an overall response rate of 67.9%, which included 91.1% of the seniors, 61.4% of the juniors and 53.3% of the sophomores at institution. The sample was 87% female and 13% male. Caucasians made up 94% of the sample and various minority groups made up the other 6%. The researchers indicated that the participants were recruited, and made no mention of how the group was selected to be the study participants.

Berg and Lindseth (2004) reported the 10 most common characteristics of effective instruction identified by participants. The most common was instructor personality, which relates to the instructor's personal behaviors and mannerism. The authors indicated that all the respondents identified instructor personality as an essential characteristic. The second was the instructors' teaching method, which encompasses the professor's apparent knowledge of subject and classroom instruction practices and whether the instruction is in line with students' educational level. Instructors' attitude, how they relate to students and their apparent concern for students was ranked third in the study. There were numerous other personal characteristics that students reported to affect the instructors' level of effectiveness and were important components of how the students rated them on SET:

- 42% of the participants indicated that how the instructor presents the material impacted effectiveness

- 33% indicated that the instructors' conduct, classroom behavior and behavior with students was important
- 33% listed content knowledge
- 28% listed their grades and the amount of work required in the course as factors in rating effective instructors.
- 25% listed the professor's availability and helpfulness
- 19% listed the instructors' enthusiasm, or interest in teaching and the course subject. (Berg & Lindseth, 2004).

Berg and Lindseth (2004) also stated that how well the instructor communicated information was one of the five most important characteristics, and fairness in the treatment of all students was the third least important. They also found that students had a tendency to enroll in courses that required the least work.

The study was limited in external validity due to the small nonrandomized convenience sample. There is a lack of generalizability due to the sample being from only one school, which has an 88% Caucasian student body and is not reflective of all other institutions. Due to the level of difference in the ratings across class ranks, the authors recommend longitudinal studies to determine if students' perception of effective instruction changes as they progress with their college education. (Berg & Lindseth, 2004).

Ratings that professors receive on SET are influenced by many factors not related to teaching, including age and years of experience (Germain & Scandura, 2005), which have contributed to the questioning of their reliability and validity. Among those is the significance that is placed on the SET (Ellis, Burke, Lomire, & McCormack, 2003) and

class size, which can affect the level of interaction between the students and the faculty (Marsh & Roche, 1997; Nerger, Viney, & Riedel II, 1997; Rodriguez-Ortiz, 1980). Due to the high emphasis on SET in faculty reviews that can affect reappointment or nonreappointment of faculty, as well as tenure and other merit decisions among others, faculty may change how they interact with students in an effort to secure higher SET ratings (Becker, 2000). A student's grade or anticipated grade is one of the factors that have been found to impact SET ratings (Baldwin, 2004; Ellis, Burke, Lomire, & McCormack, 2003; Greenwald & Gilmore, 1997; Grimes et al., 2004; Isely & Singh, 2005; Millea & Grimes, 2002; Nerger, Viney, & Riedel II, 1997). Becker (2000), and Germain and Scandura (2005) suggest that some faculty may choose to employ teaching methods that students prefer over more effective methods if this leads to instruction that is more approving to the students, even if this means that the students are not necessarily learning the most that can be learned from the class. Students' learning can be sacrificed for the sake of favorable evaluations and not be questioned because administrators consider SET and student grades in their evaluation of teaching, not what or how much the students actually learned in a course (Becker, 2000).

In his 2000 survey of the faculty of California State University, Fullerton [CSUF], Birnbaum found that faculty have acquiesced to reducing grading standards and course content in an effort to maintain their employment and secure career advancements. Birnbaum (2000) states that most faculty have discovered a way to secure higher SET ratings. The survey in which 208 CSUF faculty members participated, revealed that 65.4% of respondents believe that raising the grading standards in their courses would decrease their SET ratings, and 65.9% believed that increasing course content would have

the same effect. With this belief, 48.6% of respondents indicated that they have reduced the material presented in their courses, and 32.2% indicated that they use lower grading standards. Though they are not implementing these measures, 45.2% of the respondents believed that increasing course content would increase student learning, and 57.2% believed that raising the grading standard would have the same impact. The faculty based their responses mainly on the theory that “students will work to achieve a certain grade”. The research revealed that 70.2% of the respondents credit this decline in academic standards to the encouragement of the current system of promotion and tenure, which 92.3% of respondents stated does not give incentives for raising their grading standards (Birnbaum, 2000).

Instructors have discretion over the amount of work assigned to students and how the work is graded, but with SET there is a cost associated with this discretion. Berg and Lindseth, (2004) also found that the amount of work assigned to students and the grades they receive can have an impact on how students rate their professors on SET. In the study conducted by Berg and Lindseth (2004), 53.2% of student participants reported that the amount of work they were responsible for in a class did have an impact on how they rate the instructor in evaluations, which is inconsistent with the findings of a study conducted by Millea and Grimes (2002) revealing that “the impact of course rigor did not influence overall evaluation scores” (p. 584). Only 28% of the participants reported that their anticipated grades in a class affected how they rate the instructor on SET (Millea & Grimes, 2002). Grimes, Millea, and Woodruff (2004) found that student demographics also affected their SET ratings. They found that older students were more likely to give higher ratings on SET. Millea and Grimes (2002) report that students were more likely to

give higher ratings to faculty who are of the same gender as they. At Florida Atlantic University students in “upper division courses rated their instructors slightly higher and more effectively” (FAU, 2004, n.p.). The students’ interest in the subject matter also affect SET ratings, as this factor influences their perception of the teaching methods, previous knowledge and the knowledge gained from taking the course (Greenwald, 1997; Yunker & Yunker, 2003).

Millea and Grimes (2002) found that there also could be a difference between the impact that anticipated grades and that of actual earned grades on SET rating. The study showed that higher grades that the students have already received can increase their overall ratings of faculty on SET by 0.01, while lower grades that they expect to earn on future assignments can decrease SET ratings by 0.3. With these findings, Millea and Grimes (2002) recommend for faculty to consider “actively addressing students’ attitudes toward impending graded work” (p. 585). This action can serve as a method of generating more positive attitude and reducing the pessimism that can reduce their ratings. Having conducted this study with a sample of 149 students from Mississippi State University limits its generalizability, even when the authors have established that the institution is representative of other public universities nationwide. The authors acknowledge the need for additional research to establish methods that instructors can use to increase students’ optimism about their future work which can affect SET ratings.

Marsh and Roche (1997) propose the bias hypothesis, which states that there are other factors unrelated to the instructor’s teaching that can influence SET ratings they are assigned by students. There were several sources of bias that were found to impact the validity of SET. Those included the size of the class and the limits that places on the

possible level of interaction between students and with the instructor. A $-.02$ correlation was found between SET and average expected grades, contradicting Baldwin (2004). Greenwald and Gillmore (1997) reported a positive correlation between students' grades or anticipated grades and SET ratings. The level of correlation for other sources of bias were not provided. Similar findings also were reported by Eiszler (2002). In addition, Eiszler (2002) reported correlations between SET ratings and course popularity as well as instructor appeal. Wilson (1998) reported students' choice of course section, native language of the instructor, academic major, and student aptitude as additional factors influencing SET ratings.

Many authors, including Marsh and Roche (1997) and Greenwald and Gillmore (1997) agree that SET are not a valid methods of evaluating effectiveness of teaching. This is largely due to the numerous sources of bias that have been identified as influences on SET ratings. In some instances, students are not clearly reading the evaluation questionnaire, but instead are going down the list and bubbling uncalculated response selections (FAU, 2004).

The language used in the questionnaires is also very important and can affect students' responses on the SET (FAU, 2004). Crumbley, Henry, and Kratchman (2001) found that for nearly 50% of the participants in their study, the feeling of not having learned enough in the course was a factor that could impact SET ratings negatively. How experienced the instructor appeared to the students and the types of questions asked in class were also important factors. In the same study, being a white instructor was a factor affecting SET ratings negatively for 13.1% of the participants and for 13.3% being male had the same impact (Crumbley et al., 2001).

Marsh and Roche (1997) reported student interest in the subject, like that which is found in elective courses, lack of anonymity, instructors' presence during the evaluation were factors that tended to contribute to higher SET ratings. The value placed on the SET is a major factor, because instructors may use techniques to increase their evaluation ratings if they weigh heavily on human resource decisions (Becker, 2000; Nagle, 1998; Voge & Higbee, 2004). March and Roche (1997) contend that this factor may be contributing to the grade leniency theory. Students having admitted that the grade they expect to receive and the amount of work an instructor assigns impacts their ratings on the SET (Berg & Lindseth, 2004) is an indication that the students are not necessarily evaluating the quality of teaching, but their level of contentment with the instructor (Haskell, 1997). "As measured by coefficients that are far less than 0.7, student evaluation scores explain less than 50% of the variability and other teaching outcomes, such as test scores, scores from trained classroom observers, alumni surveys and so on" (Becker, 2000, p. 114).

Since these evaluation ratings play such a role in the human resource evaluation, instructional faculty had to adjust their teaching styles to accommodate students' wishes in order to secure the needed ratings. Instructors have resorted to lowering grading standards and decreasing workload to appeal to their student consumers (Basinger, 1997; Crumbley et al., 2001). While other concerned parties may oppose the use of SET, students support this practice which allows them to have some level of manipulation over the faculty through the students' praising or penalizing them on the SET (Crumbley et al., 2001). Crumbley et al. (2001) concur with Greenwald (1997) on the necessity for

“instruments to measure teaching performance or correct raw student ratings to remove the effects of non-instructional bias” (Crumbley et al., 2001, p. 205).

Reliability and Validity of Student Evaluations of Teaching

Marsh and Roche (1997) reviewed literature to explore the effectiveness of SET as a tool for evaluating instruction with surrounding issues of validity, utility and biases that can influence ratings. They began with the multidimensionality of teaching. Because teaching is such a versatile process and teaching styles vary, they found that the SET also needs to be multidimensional in order to capture the essence of teaching and effectively evaluate teaching effectiveness and quality. According to Arreola (2000) “there is no shortcut that will lead to a valid, fair and useful system” of evaluating faculty performance (p. xxii).

Marsh and Roche (1997) found that ratings assigned to an instructor by any two students have a correlation of .20, yet in a class of 50 students’ ratings have a correlation of .95. This number decreases to .90 in a class of 25, .74 in a class of 10 and .60 in a class of 5 students. The correlation decreased substantially as the number of students in the class decreased, reiterating the fact that SET ratings are personal and based on individual perceptions. The decrease in correlation as class size decreases, reinforces the findings that personal instructor characteristics, which are viewed and experienced differently based on class size have a significant impact on SET ratings.

Marsh and Roche (1997) reviewed a 1980 longitudinal study by Overall and Marsh, which revealed a .83 correlation between students’ initial rating of a course and a second rating of the course by the same students one year after graduation or longer. The Overall and Marsh finding shows that the raters’ personal characteristics, including state

of mind and concerns at the time of evaluation can impact their ratings. A flaw in this finding is that the authors did not state in which year the courses were taken, therefore a between evaluation time frame could not be established. The authors also concluded that there were a number of factors influencing SET ratings.

Marsh and Roche (1997) also resounded that there was no one method of measuring teaching effectiveness. As a result, they could not deem this method of using a single survey as a valid measure of evaluating teaching. Marsh and Roche (1997) found through a multisection meta-analysis that student learning was a more valid method of evaluating instruction. The authors recommend the use of a multidimensional method of evaluation that can encompass the multidimensionality of teaching.

Student Evaluations of Faculty Performance Versus Student evaluations of Teaching Strategies

Neusner (1984) presented an outline for grading professors based on the same grading criteria (A to F) that applies to students. Neusner (1984) presented specific criteria that would classify A, B, or C grade professors. Each letter was prefaced with qualification for the grade by indicating the criteria that are used for students to earn these grades. When evaluating performance, the use of SET is not appropriate as students are likely to rate instruction based on personal satisfaction factors unrelated to the effectiveness of teaching (Baldwin, 2004; Becker, 2000; Edwards, 2000; Germain & Scandura, 2005; Nagle, 1998; Voge & Higbee, 2004). Centra (1972) encouraged the inclusion of students in the evaluation of teacher performance, but only as a part of a multi-source evaluation process. The authors also suggest that students are not necessarily in a position to evaluate faculty performance because much of the faculty role

involves work outside of the classroom, of which students are not always aware.

Onwuegbuzie, Witcher, Collins, Filer, and Moore (2007) listed some of these duties which are required at different levels depending on their institutions. Yunker and Marlin (1984) suggest “achievements are relatively easy to measure and may be assessed by those personally unacquainted with the faculty member” (p. 12).

Grade Inflation

Defined

Grade inflation is defined as an increase in grades or grade point average that is not accompanied by an increase in student knowledge or student performance (Birnbaum, 1977; Boretz, 2004; Scanlan & Care, 2004; McSpirit, Kopacz, Jones, & Chapman, 2000; Nagle, 1998; Potter, Nyman & Klumpp, 2001). With grade inflation, students receive grades indicative of excellent work when their work does not meet the criteria for excellence (Nagle, 1998). Grade inflation has been found to be more prevalent in the humanities than in math and natural sciences where there are concrete answers and little room for subjectivity in grading (Wilson, 1999). This is often evaluated by comparing students’ grade point average (GPA) to standardized test scores across several decades (Birnbaum, 1977). The results of such comparisons have revealed that the students’ test scores prior to entering post secondary institutions were not increasing as rapidly as the GPAs they accumulated once at their respective institutions (McSpirit, Kopacz, Jones, & Chapman, 2000).

There have been several variations in the definition of grade inflation as this controversy broadened on college campuses throughout the country. Potter, Nyman and Klumpp (2001) affirmed that there are many different definitions or explanations of this

phenomenon. In 1983, Millman *et al.*, suggested that the value of higher letter grades (*i.e.*: A, A-) were decreasing as a result of their being so frequently used (as cited in Potter, Nyman and Klumpp, 2001). Birnbaum (1977) attributed the rise in student GPA to a reduced level of difficulty in classes making higher grades more easily attained. This trend has been referred to as “dumbing down the curriculum” (Boretz, 2004, p. 42). While all of these are part of the definition or explanation, the most commonly used definition is Lois Goldman’s 1985 definition of grade inflation: “an upward shift in the grade point average over an extended period of time” (as cited in Potter, Nyman and Klumpp, 2001, p. 9). In their 1981 article, Bejar and Blew complete this definition by adding the necessity for there to be an absence of a connected increase in educational accomplishment (as cited in Potter, Nyman and Klumpp, 2001). Birnbaum (1977) cautions the importance of ensuring that the students whose grades are being compared have comparable academic capabilities.

Theories About the Causes

Grading-leniency theory. This theory speculates that instructors would give students higher grades than their work merits in exchange for the higher SET ratings (Marsh & Roche, 1997). Voge and Higbee (2004) reported that instructors felt pressured to give students higher grades to avoid a decline in enrollment in their courses, as that is considered an indication of the quality of instruction.

The grading leniency theory is the primary theory that the literature focuses on. Gillmore and Greenwald (1999) indicate that “lenient grading” independently of quality of instruction, increases student ratings. Although Gillmore and Greenwald (1999) are familiar with the possibility that instructors would reduce grading criteria in an effort to

secure higher SET ratings, they propose an alternate explanation. Gillmore and Greenwald (1999) suggest that the change in grading and reduction of assignments may be the result of strict instructors making adjustments to their syllabi in response to student recommendations. According to Greenwald (1997), “virtually all published experimental tests” (p. 1183) of the grading leniency hypothesis were corroborated.

Student culture, pedagogical culture and institutional culture

Addy and Herring (1996) listed several factors including the withdrawal policies that allow students to drop a class when they are in danger of failing, which reduces the number of lower grades being included in student GPAs. Addy and Herring (1996) also addressed the issues of retaking a course to combine or replace a lower or failing grade with a new, higher grade. Allowing students to earn extra credit is another method that helps students boost their grades (Hassel & Lourey, 2005). Grade inflation has also been attributed to the need to protect students’ self-concept (Edwards, 2000). Rather than allowing students to fail, measures have been put in place to shelter them from hardships, and keep them from feeling failure in order to prevent the lowered self-esteem that can accompany failing (Edwards, 2000). This method of sheltering students from risks that could afford them the opportunity to realize personal successes and accomplishment (Edwards, 2000) is denying students their need to earn their self esteem (Edwards, 2000) and attain “self-actualization” (Zastrow & Kirst-Ashman, 2001, p. 433). Edwards (2000) proposed that this may be in part why there has been an increase in depression and the earlier ages of onset.

Birnbaum (1977) conducted a study at the University of Wisconsin, Oshkosh to test hypotheses on grade inflation. Birnbaum (1977) used existing data collected at the

university from 1968 to 1969 and from 1974 to 1975. The information included grade point average (GPA) and scores from the Undergraduate Record Exam (URE) from 1968 to 1969 and the Graduate Record Exam (GRE) from 1974 to 1975. The sample included all seniors at the university at the time of data collection. Students for whom “certain information” (Birnbaum, 1977, p. 529) was not available were excluded. After exclusions, the earlier group had 769 valid cases, and the latter had 535 cases. The author did not clarify what “certain information” constituted. However, the most common reason for exclusion was that the students’ rank-in-class was not available to the researcher.

The hypotheses to be tested included the following reasons for increase in GPA: 1) brighter students were coming to the university, 2) students’ learning increased with student achievement, 3) higher number of female students enrolled, 4) students matriculating in fields where grades generally are higher increased, while enrollment in fields where grades have been lower decreased, 5) students taking different courses, and 6) increased use of pass or fail grades and withdrawal options.

In his study of the two groups, Birnbaum (1977) observed that the students’ high school rank-in-class increased from 68th to 71st percentile as they entered the university. The average first semester junior year GPA increased from 2.68 to 3.17 which was not congruent with GRE scores decreasing from 492 to 477. Birnbaum (1977) noted that these differences were consistent with national trends. In the first group, students majoring in education, music and physical education accounted for 24% of the sample, and students majoring in sociology, English, history, geography, political science and economics accounted for 37%. In the second group, they represented 46% and 17%,

respectively. Female students increase from 40% in the first group to 48% in the second group. Birnbaum (1977) used these factors to support the theory that what is believed to be grade inflation is the result of a change in students and their abilities, rather than a change in grading practices.

Regression equations were used to test the hypothesis of a difference in the quality of students coming to the university using gender and major subgroups as catalysts. The students' class rank in high school was the independent variable and their university GPA was the dependent variable. Based on the information for the group from 1968 to 1969 a predicted GPA of 2.65 was projected, which was lower than the actual 2.68 initial average and 3.17 in the latter years. This projection was found to be "within 0.1 grade points of actual 1968-69 averages in 21 of the 25 subgroups and at least 0.3 grade points lower than actual 1974-75 averages in 19 of the 25 subgroups" (Birnbaum, 1977, p. 531). The test did not support the hypothesis that the higher grades are indicative of a change in students rather than a change in grading criteria.

The second hypothesis that improvement in student achievement was the reason for the increase of student GPAs were also tested with regression equations for the 25 major subgroups. The independent variables in these equations were the URE and GRE scores and the students' university GPA, the dependent variables. The projected GPA of 2.63 yielded similar issues as the test of the first hypothesis and also did not support this hypothesis. The hypothesis that grade inflation was a result of a change in gender and major distributions was tested by calculating university mean GPA. The result was a 2.68 projected GPA which was the same as the 1968-69 actual GPA and therefore failed to support the hypothesis. The final hypothesis that changes in students' selection of

courses resulted in grade inflation also was not supported. The researcher found that 40% of the courses taken by the first group were different from those taken by the second group. Using the remaining courses that both groups had taken, the comparison did not yield statistically significant results. There was some support for the hypothesis that a change in grading policies may have resulted in the increase in GPA between the groups. By recalculating P and W grades as C and D grades, the researcher found results to indicate that if students were not able to opt for a P grade or withdraw from classes they are not doing well in, the average student GPA would be lower.

The study focused on students from one university which historically attracted students in the 70th percentile of their graduating high school classes. The results were not compared to that of other schools. Although Birnbaum (1977) reported that demographic information was gathered, this information was not reported with the findings. Generalizability is limited to the school. The study also was limited to the analysis of data that could be obtained from existing university records. By including other institutions, the researcher could have increased the generalizability of this research. By including more than just two academic years in the study, the researcher could have improved validity of the results by showing more than just one set of data reflecting the differences. Falkenberg (1996) cited those similar potential causes of grade inflation.

While the threat of being drafted has subsided, students are faced with new threats that create a need to maintain a certain GPA. One of the most popular is the need to maintain a minimum required GPA for continued enrollment in some programs or even a graduation requirement (Addy & Herring, 1996). The lack of consistency in grading

procedures is also a factor that may be contributing to grade inflation, as different faculty members may have different grading standards, where some may have lower performance requirements to reach “A” quality (Addy & Herring, 1996; Boretz, 2004; Nagle, 1998). This lack of consistency has also been attributed to a “degradation of academic standards” (Birnbaum, 1977, p. 522). One of the most controversial threats lays in the notion that faculty would grade students more easily to get higher ratings on student evaluations of teaching or to attract students to register for their classes (Addy & Herring, 1996; Voge & Higbee, 2004).

With the uses of SET moving from a tool to help instructors improve their teaching to becoming a tool used in making human resource decisions, the pressure to give higher grades has increased significantly (Algozinne et al., 2004; Eiszler, 2002; Parayitam, Desai & Phelps, 2007). Faculty members who are not tenured have been found to assign higher grades than their tenured colleagues (Boretz, 2004). There is also great pressure on schools to please their consumers [students] and give them the goods [grades] they are paying for (Hassel & Lourey, 2005), especially when the students value their grades more than the education that they are receiving (Hassel & Lourey, 2005; Isely & Singh).

Impact (negative) of Grade Inflation on Students and Institutional Accountability

Grade inflation places students at a higher level of perceived performance than they ordinarily would have attained. When employers or graduate programs screen students, they see a group of high performers, when in actuality not all of them are performing as well (Anonymous, 2002). The grades students are receiving lack the “observable evidence of learning” that Gray (2002, p. 53) states should be a part of

assessments. While they are receiving higher grades, many students come prepared to do as little work as they can to get by (Basinger, 1997). It has been cited that many students come into higher education intending to invest minimal time and effort in courses while receiving high grades (Basinger, 1997). Students become accustomed to receiving grades that they are not earning but have paid for in tuition (Hassel & Lourey, 2005). Higher grades lose the value and respect that were once attributed to them, and institutions of higher education also lose credibility in their evaluation of students. Students progressively lose their sense of personal responsibility and accountability for their education and their grades (Hassell & Lourey, 2005).

Educators' expectations for students' learning are lower, as a result, students are being sent into the workforce or graduate programs ill-prepared for the demands that will be placed upon them. Students are not challenged or pushed to achieve at their highest potential and they spend their time in college barely getting by, but still making the grade due to lowered standards (Edwards, 2000). Hassel and Lourey (2005) provide information to suggest that students learn to expect others to take responsibility for them. Rather than seeing their failing grades as faults needing to be remedied, many students see them as an indication that extra credit is needed in the course to improve their grades (Hassel & Lourey, 2005). In their study of college students, Hassel and Lourey (2005) found that "62% think it is an instructor's responsibility to offer extra credit, and 52% expect instructors to be flexible in grading" (p. 7).

Grade Inflation and Student Evaluations of Teaching (Studies)

Ellis, Burke, Lomire, and McCormack (2003), conducted a quantitative study to verify that there was a relationship between student grades and student ratings on SET.

The study consisted of a non-probability sample of 5602 students enrolled in 24 instructors' courses at Minot University. Students in all the courses had to complete SET, although the university only requires non-tenured instructors to have SET in all their courses, while tenured faculty only had SET in two of their classes per year. In the final week of class, the students were asked to evaluate the course and the instructor on separate questionnaires containing items rated on a 10 point scale. The researchers completed correlation analyses to compute the data using the class GPA as the independent variable and course and instructor ratings as dependent variables. They found significant correlation ($p < .01$) between the class GPA and instructor ratings. When they completed a multiple regression analysis in controlling the class GPA, they found that there was no longer statistical significance.

Professors who are known to give higher grades are likely to have more favorable student evaluations (Voge & Higbee, 2004). Researchers have found that there is a positive correlation between students' grades and the instructors' SET ratings (Baldwin, 2004; Ellis, Burke, Lomire, & McCormack, 2003; Isely & Singh, 2005). There are many possible explanations for this factor. One may be that the professors are more effective and students in their classes are indeed learning more and earning higher grades (Ellis, Burke, Lomire, & McCormack, 2003; Greenwald & Gillmore, 1997). Another is that the professors know that those who give higher grades receive more favorable SET ratings, so they give the grades that get them the desired ratings (Ellis, Burke, Lomire, & McCormack). McSpirit et al. (2000) found that "three-quarters of non-tenured faculty concerned over grades and tenure concede being influenced by student ratings of their performance when it becomes time to grade student performance" (p. 24). In turn, the

students show their appreciation through the higher SET ratings (Greenwald & Gillmore, 1997). Greenwald and Gillmore (1997) suggest that giving higher grades is one of the more uncomplicated ways for instructors to increase their SET ratings. Instructors can also design their classes to appeal to students by reducing the workload or changing certain criteria (Armstrong, 1998).

Dresner (2005) proposes a unified grading theory. This would require that standard principles be developed to provide a basis for grading student work. With this method, faculty can move away from relative grading, which includes factors other than academic performance. The practice of absolute grading could then be implemented, where faculty would grade student work based on predetermined criteria. This process would provide a concrete basis for grading that would reduce grading based on personal aspects or personality and potentially reduce the incidence of grade inflation.

Trends, Controversy, and Emerging Strategies to Manage Grade Inflation

The grade inflation controversy has been on the rise for the past four decades (Wilson, 1999; Potter, Nyman and Klumpp, 2001). Grades, which used to be a reflection of students' academic performance, are now being viewed as a product that the students, the customers, pay for. As a result, faculty are giving grades to gratify and appease these customers (Basinger, 1997). In the 1960s, when the problem with grade inflation was developing, it was in an effort to assist draft eligible students to maintain their draft deferments (Voge & Higbee, 2004; Anonymous, 2002; Birnbaum, 1977). Today, grade inflation, is in part, a result of the rise in student consumerism, which leads to the assigning of grades that are worth the money the students pay to be in school (Anonymous, 2002, Hassel & Lourey, 2005; Kanagaretnam & Thevaranjan, n.d.;

Scanlan, 2004). Some faculty also have taken the potential impact of a lower grade on a student's self-esteem into account in their grading, and this has in part resulted in the students learning to expect the higher grades without performing the work to earn them (Boretz, 2004).

It has been observed that college student grades throughout the country are rising significantly (Addy & Herring, 1996; Anonymous, 2002; Birnbaum, 1977; Boretz, 2004; Edwards, 2000; Ellis, Burke, Lomire, & McCormack, 2003; Kohn, 2002). The number of students receiving an A- or higher at many schools has increased significantly from where it was 20 years ago and the number of Cs has decreased (Edwards, 2000; Ellis, Burke, Lomire, & McCormack; Kohn, 2000). Close to 50% of students receive grades of A- or higher (Johnson, 2003). At the same time, the number of students receiving lower grades such as a C+ or lower has decreased to fewer than 20% (Johnson, 2003). While many believe that the lower standards and the increase in the number of students receiving unearned grades resulted in grade inflation, others are open to the possibility that academic performance may warrant the increase in grades (Edwards, 2000; Johnson, 2003).

Harvard University, which has been one of the institutions at the center of this controversy, solicited its undergraduate faculty's assistance to evaluate their grading practices as they begin defining grading standards and developing uniform grading practices (Anonymous, 2002). Grade rationing as it is being implemented at Princeton University (Roarty, 2004, Voge & Higbee, 2004), has been a method used in effort to combat the issue of grade inflation. While this matter is being addressed at institutions throughout the country, few solutions are being implemented. Armstrong (1998)

proposes direct assessment of students using pre- and post-tests as a measure of evaluating teaching effectiveness and gauging learning [value added]. The use of outside evaluators would persuade faculty to use methods to increase classroom learning (Armstrong, 1998).

Rather than attempting to fight grade inflation, changing methods of teaching to increase knowledge to commensurate with the higher grades is another option for faculty (Basinger, 1997). Another method is to have students complete evaluations of the course and instructor separately (Ellis, Burke, Lomire, & McCormack, 2003). Nevertheless, Ellis, Burke, Lomire, and McCormack (2003) recommend that students' evaluations be used in conjunction with other methods. Kanagaretnam, Mathieu and Thevaranjan (2003) warn that placing "excessive weight on student evaluations can have negative consequence" (p. 7). They listed grade inflation, including lowered grading standards and decreased student effort, among the potential consequences. Nagle (1998) presented the example of how two universities are attempting to address the problem:

Duke University recently proposed to replace its standard grading system with a new measure called the "achievement index." The index takes into consideration the range of grade distributions in a particular class. Student QPAs (Quality Point Average) are adjusted upward or downward based on the level of difficulty of the course as indicated by the class's grade distribution. Use of the index was eventually rejected by a Duke faculty committee. (p. 41)

Indiana University has expanded the information provided on each student's transcript. In addition to the student's grade for a particular class, information is provided about the grade distribution, median grade, student rank, and average

student QPA for each class. Though these data provide insights into a student's relative performance, they are hardly concise. This procedure also places an additional onus on faculty to undertake the cumbersome task of ranking, from top to bottom, each student in the class. (p. 41)

Faculty Assessments, Promotion and Tenure, (FAPT) Decisions

Institutional Decisions on Faculty Assessments, Promotion and Tenure

The literature indicates that there has been an increasingly heavy reliance on the instructor's SET ratings for the purpose of making faculty assessments, promotion and tenure, decisions (Centra & Gaubatz, 2002; Ellis, Burke, Lomire, & McCormack, 2003; McAlpine & Harris, 2002; Nerger, Viney, & Riedel II, 1997). Teaching is the primary role of many faculty in higher education, which is often evaluated through the use of SET. The Carnegie Foundation found that, "teaching has a 50 to 60 percent weight in personnel decisions" (Becker, 2000, p. 113). In research universities, teaching has a 25 to 30% weight in these decisions and "unacceptable teaching" (p.113) has prevented researchers from getting tenure or promotions (Becker, 2000). Unfortunately, the literature does not provide a clear history of when the SET started to become such an integral part of the institutional evaluation of faculty, or how much weight is generally placed on them.

Influence of Student Evaluations of Teaching in FAPT Decisions

Institutions are increasingly using SET "to determine worthiness for merit pay, retention, tenure and promotions" (Berg & Lindseth, 2004, p. 565). Because of this, professors may give higher grades to receive favorable SET ratings (Birnbaum, 1977; Crumbley et. al., 2001; McSpirit et. al., 2000). Basinger (1997) suggested that these uses

of SET are a method being implemented to discourage faculty from placing higher standards on students that would lead to greater demands from the students and make higher grades more difficult to attain. Such a situation would reduce an institution's competitive advantage if other institutions are not implementing the same practices (Basinger, 1997). Arreola (2000) proposed the alternative reasoning that the increased value placed on the SET may be a result of the higher demand for accountability from institutions of higher education. In the literature reviewed, the authors did not provide very specific information on how much FAPT decisions are impacted by SET. However, the authors affirmed the SET as an important part in the decision process and that SET weights heavily (Baldwin, 2004; Isely & Singh, 2005; McSpirit et. al., 2000; Nagle, 1998).

Through this literature review there was a noticeable lack of research exploring the faculty perception. Much of the research that has been conducted involved comparison of student grades over time as researchers investigated grade inflation. While researchers postulated that the use of SET in FAPT decisions as one of the factors contributing to grade inflation, research investigating this phenomenon is not available. Voge & Higbee, (2004) did include some direct faculty quotations and other accounts in their article, however their interviews were very limited, which limits generalizability to the faculty population. After reviewing the available literature, the need for studies exploring the faculty perception of the use of SET in FAPT decision is apparent. There is also a need for studies exploring faculty's perception of reports that faculty are inflating student grades to secure higher SET ratings. The proposed study will contribute to the literature by addressing these identified gaps.

Discussion of the Literature

Theoretical Literature

Construct of teaching effectiveness: Models of student evaluations of teaching.

According to the prevailing literature, there is no one method that can fully capture the essence of teaching (Marsh & Roche, 1997). Teaching effectiveness involves many different factors including instructor personality and student needs. The effective methods of teaching for one instructor in one class may not have the same results in a different classroom. Different schools have been working to develop more effective methods of evaluating classroom teaching. The recommendation is to use more than one method to perform this challenging task (Centra, 1972; Marsh & Roche, 1997).

Grade inflation. Nationwide, grade inflation is believed to be plaguing the higher education system. Researchers report that the causes of grade inflation have progressed throughout the past five decades from use in assisting students in ensuring enrolment with good academic standing in higher education, to avoid being drafted into the armed forces in the sixties, (Birnbaum, 1977; Voge & Higbee, 2004) to financial concerns in this decade. The increasing use of SET in FAPT decisions and the rise of student consumerism are two of the most commonly named culprits. First is the belief that instructors who give higher grades receive higher SET ratings, which they need for FAPT decisions. Next is the need to give the students value for their money. Since grades are seen as a measure of the education, the higher the grade, the higher the perceived quality of the education received. The primary concern with grade inflation is the notion that it compromises the quality of higher education by classifying students as A students who

are not producing A quality work. Students end up receiving high quality grades but the education received is not congruent (Edwards, 2000).

Institutional decisions on faculty reappointment, promotion and tenure. While authors postulate that the use of SET in human resources decisions is having an adverse effect on the quality of higher education, the researchers did not confirm the extent to which this is true. Through this literature review, only one author (Becker, 2000) presented potential weight of the SET on FAPT decisions. The rest of the literature simply mentions this concept with no elaboration or exploration.

Empirical Literature

Reliability and validity of student evaluations of teaching. The review of the literature revealed that there has been significant correlations noted between SET and student grades received or anticipated (Baldwin, 2004; Becker, 2000; Edwards, 2000; Nagle, 1998; Voge & Higbee, 2004). Studies have been conducted to explore the correlation between SET and grade inflation, which were found to be statistically significant (Armstrong, 1998; Centra & Gaubatz, 2002; Greenwald & Gillmore, 1997; Millea & Grimes, 2002). The professors interviewed in Voge and Higbee's (2004) article admitted that university professors are indeed giving students higher grades than deserved out of concern for the potential repercussions of assigning students the lower grades actually earned. This is repeated throughout the literature. Researchers have also raised concerns that students are not educated on how to effectively complete the evaluation (Becker, 2000; Harvey, 1999).

Grade inflation, student evaluations of teaching, and FAPT decisions.

Throughout the literature the use of SET in FAPT has been presented as one of the many

reasons that grade inflation is so prevalent in higher education (Baldwin, 2004; Ellis, Burke, Lomire, & McCormack, 2003; Germain and Scandura, 2005; Greenwald & Gilmore, 1997; Grimes et al., 2004; Becker; 2000; Isely & Singh, 2005; Millea & Grimes, 2002). The literature has revealed that there is a significant correlation between SET and student grades. The researchers have not confirmed that the rise in grades is not a result of improvement in student learning and performance, but there is considerable information to suggest that there is a positive correlation (Baldwin, 2004; Ellis, Burke, Lomire, & McCormack, 2003; Greenwald & Gillmore, 1997).

No studies were found to determine the extent of the influence of SET ratings on FAPT decisions. There is also a lack of research to provide information to support the notion that faculty members are inflating student grades or lowering standards to secure higher SET ratings.

Conclusions

The constant theme in the literature is that the use of SET in FAPT decisions has contributed to a factor that may be influencing a downward cycle in the quality of higher education. The research indicates that there are many extraneous factors influencing SET ratings that are not particularly related to teaching, including faculty's assignment of higher grades to secure higher SET ratings. The influence of these factors has therefore compromised the validity of this process of evaluating teaching since higher grades are not indicative of increased learning.

This lack of validity is a result of the lack of training of students who are asked to be observers and complete faculty evaluations. Such training is necessary to lend validity and reliability to the evaluation (AAUP, 2001). The student evaluations of teaching have been biased in that students rate professors based on outside factors such as anticipated grades rather than the methods and effectiveness of instruction. The research implies that

the value placed on SET in the FAPT decision has led faculty to use creative grading to secure higher SET ratings. Instructors have acquiesced to lowering standards and giving underserved grades in order to get favorable evaluations because their continued relationship with their institutions and the quality of these relationships are contingent on these evaluations (Voge & Higbee, 2004).

At this time a viable solution has not been reached, but individual institutions are making attempts to address this issue through policies such as grade-rationing, which is being implemented at Princeton University (Roarty, 2004).

CHAPTER III

Research Design and Procedure

Research Methodology

This was a non-experimental, quantitative, exploratory study aimed at better understanding faculty perceptions of the role that SET play in grading and FAPT decisions. This task was accomplished by exploring the faculty perceptions of the uses of student evaluations of teaching in faculty assessments, promotion and tenure decisions and its impact on grade distribution through an anonymous online survey.

Research Design

Permission was obtained from the Institutional Review Board at Lynn University to conduct this study using an online survey. The qualifying population consisted of current ranked faculty at the eleven institutions in Florida's State University System (SUS) and 28 Independent Colleges and Universities of Florida (ICUF) member institutions from which a non-probability criterion sample was derived. All 39 institutions have Level II or higher accreditation by the Southern Association of Colleges and Schools (SACS). The survey was made available to the qualifying population for a period of four weeks from the date it was released online. The selected non-probability criterion sample was invited through an email to participate in this study by completing the online survey. A link and the password to access the survey were provided in the message. Invited candidates were able to follow the link to the consent page which reiterated that participation in the survey was anonymous, and advised them that clicking the provided "I agree" button at the bottom of the form constitutes consent to participate

in this research. Respondents who acknowledged the statement by checking the provided box, were taken directly to the 36 question survey which they were able to complete anonymously.

Population and Sample

The population consisted of faculty at Florida's institutions of higher education. Participation in this study was limited to faculty who are employed in Florida and are ranked as Instructors, Assistant Professors, Associate Professors, or Professors. The target population consisted of faculty teaching at eleven SUS institutions and 28 ICUF member institutions and hold a rank as Instructor, Assistant Professor, Associate Professor or Professor. Faculty Demographics (2000) show that there are 14,177 full-time faculty employed in public institutions, and 4,633 in private institutions. This report did not indicate how many of the public institution faculty were employed at the 11 SUS faculty. There was also no distinction to indicate the number of ICUF faculty.

A power analysis was conducted using the G* Power software (Faul, Erdfelder, Lang, & Buchner, 2007), which revealed that using the total population of 18,810, a total sample size of 319 participants was necessary to be able to analyze the data with effect size (the weight of the impact of the independent variable on the dependent variable) of .20, power level (the probability that findings are correct) of .95 and error probability (the chances of study findings being a result of sampling error) at .05 (Rubin & Babbie, 2001). Using Israel's (2003) Simplified Formula For Proportions a necessary sample size of 392 participants was found The invitation to participate was sent to 6,400 faculty members from 26 of the selected schools, using a combination of direct emails and distribution through the institutions.

The response rate of the survey impacts the size of qualified data set available for analysis. The higher the response rate, the more likely it will be that the data producing sample is representative of the population, which would reduce the probability of findings being the result of error (Rubin & Babbie 2001). The sample's representation of the population is vital to generalizability of the study's findings. Roy and Berger (2005) found an 8.42% return rate on in email surveys such as the one used in this study. They also noted a difference in return rates from the different countries where participants were invited to complete the survey, and the return rate from the USA was 10.36%. Dommeyer and Moriarty (2000) found a similar return rate of 8% for this type of survey. Dommeyer and Moriarty also noted the increased ease of discarding the email invitations or just deleting them with a single mouse click, which may be a contributing factor to the low response rate for this survey format. These response rates are slightly higher than the response rate necessary to attain the 262 responses needed in this study. The researchers suggest the method of acquiring email addresses used to invite participants to the survey as factors which could influence response rate. Therefore, receiving the email addresses for this survey directly from participants' schools may positively impact the response rate.

Evaluation of Sampling Design

In an effort to increase sample reliability, a large population was selected from which to derive a sample. Exclusions that were made, such as the rank requirements, insured more uniformity in the sample. A ranked instructor for example is expected to spend more time teaching than an adjunct who may be teaching one class. Adjunct faculty were excluded since their level of exposure to students, SET and its uses may

vary from that of the targeted population. The number of years of experiences between the two may also differ as the ranked faculty are more likely to have had some experience as an adjunct prior to securing a full-time position. The sample is only representative of ICUF and SUS faculty. It may share some characteristics with other institutions in Florida, and throughout academia, to the extent that the 11 SUS and 28 ICUF institutions are representative of these other institutions.

The survey was anonymously administered on SurveyMonkey.com. Anonymity was maintained to the extent provided by law and the technology used. Before the survey was completed, participants were advised of the voluntary nature of participation and their right to decline to participate, which they acknowledged in the consent form. Although stated in the invitation email, the consent form also reiterated to participants that they were consenting to participate in the study by completing the survey, and that the survey's anonymous format limits the researcher's ability to honor revocations of consent as individual responses are not identifiable.

In an effort to maintain the anonymity of the participants and employers, the individual schools at which faculty were invited to participate in the study are not identified. The participating institutions are identified as Independent Colleges and Universities of Florida (ICUF) and State University System (SUS). The survey site was also set to not keep participant IP addresses in an effort to further protect anonymity.

Data Collection Procedure

The purpose of all data collection was to explore faculty perception of the use of SET in FAPT decisions and the impact this process has on overall teaching and grade distribution. Data was collected using an online survey of qualifying faculty. The 36

questions on the survey were structured within four (4) sections: demographics, institution information, faculty's perception of the use of Student Evaluations of Teaching including FAPT decisions, and the degree to which faculty believe that Student Evaluations of Teaching impact grade distribution. Only those invited to participate in the survey, who received the password were able to follow a link in the email to complete.

Instruments

Data was collected using the 36 question survey designed by the researcher for the purpose of this study. This questionnaire was an original instrument, designed specifically to probe into the questions which the researcher sought to investigate. A factor analysis was used to complete an initial assessment of validity of each subscale and the total survey. The survey was pilot tested with faculty from Lynn University after IRB approval had been granted. The purpose of the pilot testing was to assess the validity of the survey so that improvements could be made if necessary. The pilot was used to ensure that survey questions were clear to participants, and elicit the responses researcher intended to explore. After completing the pilot testing and making the necessary changes, the survey was made available online where it was available for four weeks. Data was collected through the SurveyMonkey.com website. The digital copies of the survey responses are maintained by the researcher and available for review by researcher as needed. Data sets were assigned a chronological identification number and IP addresses were not maintained or included with data to ensure anonymity.

Data Analysis

The data analysis began after the four-week period of data collection. The data collected from the surveys were entered in a Statistical Package for the Social Sciences (SPSS) database for analysis, which was the software used to compute data. Data entry was completed using the SurveyMonkey service, which allowed the researcher to retrieve the electronic data already in a spreadsheet where the information could be copied into an SPSS spreadsheet for analysis. The analysis included descriptive statistics, bivariate correlation analyses, t-test, and Chi-square.

The first analysis was to complete descriptive statistics to explore the demographics of the respondents to be compared to demographic information reported in United States (2000) to determine if the sample is representative of the population. The rate at which professors are in agreement or disagreement in their responses to the questions on the survey also were explored. The data was analyzed to determine if professors were reporting a belief that SET ratings impact faculty's grading practices, and well as the percentage reporting having considered their ratings as they prepared assignments and assigned grades.

Bivariate correlation analyses were conducted to determine the relationship between size of the student population, and classroom size, and belief that SET ratings are considered in the preparation of assignments, as well as consideration in grading and for reports of engaging in these behaviors themselves. Chi-square analysis was conducted to compare school setting (*i.e.*: SUS or ICUF) and belief that SET ratings are considered in the preparation of assignments, as well as consideration in grading and for

reports of engaging in this behavior. T-tests were used to compare school setting to reported impact of SET on FAPT decisions.

Chi-square analysis was conducted to assess the relationship between the occurrence of the belief and implementation of the behavior. Faculty's reports of considering SET ratings in their preparation of assignments and grading was assessed using similar correlation analyses. Relationships between faculty responses and specific characteristics such as years in the field and tenure status were also explored using correlation analyses. Bivariate correlations were used to assess for a relationship between the availability of tenure and the beliefs that faculty consider SET ratings in preparing assignments and/or grading. Correlations analyses between the beliefs that faculty consider SET ratings in preparing assignments and/or grading and the participants' tenure status where it is offered, were also analyzed. The analysis was repeated using school setting as the grouping variable.

Bivariate correlation analyses were conducted between all variables. Specific areas of interest were to determine if there is a relationship between the years of experience and belief that faculty consider SET ratings in grading and assignment preparations, report of doing that, and their interest in a unified grading system. Multivariate analyses were also used to explore co-occurrences. Correlations and Chi-square analyses were completed to explore the impact that sets of factors including school setting and professor characteristics have on reports or beliefs of the existence of grade inflation due to SET rating. An analysis was conducted to explore the commonalities in faculty who report having considered SET ratings in their own assignment of grades as well as those who have done so in development of student

assignments. Multivariate analyses were used to identify the similar characteristics of the faculty who did self-select to participate in the study. Similar analyses were also used to explore differences in faculty who believe SET is impacting grade inflation as well as those who report having considered their ratings when grading students' performances.

Chi-square was used to compare responses of SUS faculty to ICUF faculty on categorical variables, T-tests were performed to compare means of continuous variables including:

- Reported weight for SET on FAPT decisions by Administrators and that perceived by faculty
- Perception of SET impact on grade distribution and interest in a unified grading system
- Faculty perception that SET ratings are motivated by grades and agreement with use of SET in FAPT decisions.

Hypothesis testing

Hypothesis 1: ICUF faculty will experience greater pressure to have high SET ratings than SUS faculty. Chi-square analyses were conducted to explore the differences between ICUF and SUS faculty responses.

The faculty's responses to question 18 and 21 were compared according to school types (SUS and ICUF) and availability of tenure. Chi-square analyses were used to assess whether the number of respondents from ICUF schools reporting having been advised that SET will impact their annual evaluations and those who have been advised that SET ratings impact promotion and tenure decisions is greater than the number of

respondents from SUS schools. Chi-square analyses were conducted on section 3 questions 18 and 21, with question 8 as the grouping variable.

T-tests were conducted to explore whether or not faculty at ICUF institutions reported that SET ratings having greater value on their annual evaluations and overall FAPT decisions compared to SUS faculty. The groups being compared in this analysis are SUS and ICUF institutions. The tests were repeated to compare faculty according to tenure status. The variables being tested were the percentage values they report are assigned to SET in FAPT decisions. Section 3 questions 20 and 22 were used to conduct these tests, with question 8 and 14 as the grouping variable.

Hypothesis 2: SUS faculty will report less impact of SET on their grading practices than their ICUF counterparts, SUS faculty will report less impact of SET on their grading practices than their ICUF counterparts was tested using Chi-square test for independence. Chi-square analyses were conducted using item responses from section 3 question 8 with section 4 questions 23-29. The analyses were repeated using tenure status as the grouping variable. Correlation analyses between these variables were also performed.

Hypothesis 3: Class size is inversely related to perceived importance of SET ratings. Bivariate correlations were used to assess whether less importance is placed on SET in schools with larger populations and larger classes. The survey items used in these tests were section 3 questions 10 and 20, section 3 questions 10 and 22, section 3 questions 11 and 20 and section 3 questions 11 and 22.

Hypothesis 4: Faculty who perceive the importance of SET on FAPT to be high are more likely to consider SET ratings in preparing student assignments and grading.

Bivariate correlation matrix between item responses from section 3 question 20 and 22 and Section 4 question 23-29 of the questionnaire were used to test this hypothesis.

CHAPTER IV

RESULTS

The purpose of this research was to explore faculty perceptions of the use of SET ratings in FAPT decisions. The study investigated whether faculty perceived a correlation between SET and grade distribution. A 36 question survey instrument was developed specifically for this study. The survey was pilot tested for validity with Lynn University faculty. The pilot study was completed over a four day period, and 22 responses were received when the pilot test concluded. Only minor editing recommendations were made. They were completed before the survey was distributed to the target population.

The 6,400 individual who were invited to participate include 5,700 members of the faculty at 17 institutions to whom the invitation to participate in this survey was sent directly. There were eight institutions which distributed the survey invitation internally from their distribution lists. There was one institution where only some departments agreed to share the link with faculty to participate in the survey. The number of faculty who received the invitation through the latter two methods was approximately 700, as reported by the individuals who received and distributed the emails at each school. One week prior to the end of the data collection period, an email reminder was sent to potential participants. At the end of the four week data collection period the survey received a 12% response rate, from which 711 participants completed the survey. There were 50 respondents who did not indicate that they were ranked faculty, including adjuncts and other faculty who hold more than one position (*i.e.*: Department Chairs,

Deans). Based on their description of “other” in their responses to the rank question, they were recoded to be included in the sample when applicable. There were a small number of adjunct faculty who responded, but the number was too small for significant analysis and were therefore excluded from the sample along with any other responses that could not be recoded into a rank category. The final number of qualifying responses available for analysis was 704. The sample size used in this study is 180% of the minimum sample size found using Israel’s (2003) formula, and 221% of the required sample size from Faul, Erdfelder, Lang, and Buchner’s (2007) G*Power sample size calculator.

Descriptive statistics were used to examine the sample’s demographic information and explore the sample’s representativeness of the larger population. The demographic information for this sample is illustrated in Table 1. In comparison to United States census data from 2000, the African American and Asian faculty were under represented in this sample. The representation of both groups in this sample was half of their representation in the population. The White faculty were over represented by more than 8%. There was not a specific category addressing ethnicity, though several Hispanic faculty noted that they selected the White category. The representation of American Indian or Alaska Native faculty was accurate. The census data did not provide information for Native Hawaiian or Other Pacific Islander faculty, therefore a comparison could not be made. (United States, 2000).

The participants were on average 50 years old, though the national average was 44 years old. They had an average of 14 years of experience in post secondary education. Nationally, one third of faculty hold a doctorate degree and the highest grade held by another third is a masters degree. The highest level of education completed by 84% of

the respondents was a doctorate degree. The sample included an even distribution of Professors, Associate Professors and Assistant professors, with nearly 30% of the sample from each group. This number is representative of the Florida faculty according to survey information collected from the institutions by Integrated Postsecondary Education Data System (IPEDS). Approximately 44% of participants reported that their annual salaries from teaching range between \$50,000 and \$75,000, which is more than twice the percentage reported in the US Census. (United States, 2000)

Though 67% of the respondents were SUS faculty and 33% were ICUF faculty, these proportions are representative of the population as the SUS, with a significantly larger student population, does employ more faculty than the ICUF. The ratio of SUS to ICUF faculty is nearly 3:1, although the faculty population at five schools was not available. With the faculty population report for these five schools added, the ratio may be closer to the 2:1 ratio in this study. The significant difference in educational attainment as well as income which was found between the two groups may be a result of the inclusion of Community College faculty, Lecturers as well and other factors which were not excluded. The census data included the national faculty population. The data also did not make the exclusions that were made in this study. While there are some significant differences noted between the sample and the faculty population they may be explained by the exclusions made in the sample. The comparison between the sample and the census data is shown in Table 1, though some census information was missing.

The questions in the last three sections of the survey instrument were designed specifically to help answer the research questions and test the hypotheses in this study, and were used to conduct the following analyses.

Table 1

Demographic Characteristics

	Total sample	U.S. Census	Group Differences χ^2	SUS	ICUF	Group Differences χ^2
Gender			.98			.97
Male	54.1%	56%		53.7%	55.0%	
Female	45.9%	44%		46.3%	45.0%	
Race			32.28***			1.14
American Indian or Alaska Native	.4%	1% ^{††}		.4%	.4%	
Asian	2.2%	6% ^{††}		2.6%	1.3%	
Black or African American	4.1%	8% ^{††}		4.1%	4.0%	
Native Hawaiian or Other Pacific Islander	.1%	.03% ^{††}		92.9%	94.2%	
White	93.3%	85% ^{††}				
Education Level			747.11***			46.46***
Bachelor	1.0%	22%		.8%	1.3%	
Master	14.5%	29%		8.2%	27.4%	
Doctorate	84.5%	29%		90.9%	71.3%	
Institutional Rank			2.71			10.98*
Instructor	9.4%	9.9% [†]		8.5%	11.3%	
Assistant Professor	32.5%	28.4% [†]		29.9%	37.8%	
Associate Professor	29.8%	27% [†]		33.5%	22.2%	
Professor	28.3%	27.8% [†]		28.2%	28.7%	
Annual Teaching Income			1155.14***			109.19***
Less than \$30,000	2.0%	45%		1.3%	3.5%	
\$30,000 - \$50,000	15.5%	26%		7.0%	32.8%	
\$50,000 - \$75,000	42.3%	19%		41.1%	45.0%	
More than \$75,000	40.2%	10%		50.6%	18.8%	
Tenure Status						.89
Not on a tenure track	30.7%			17.1%	59.5%	
On a tenure track	25.8%			28.8%	19.4%	
Tenured	43.6%			54.2%	21.2%	
SET Form Used						1.12
No	1.0%			1.3%	.4%	
Yes	99.0%			98.7%	99.6%	
	Sample Mean	U.S. Census	<i>t</i>	ICUF Mean	SUS Mean	<i>t</i>
Age	50	44	13***	50	50	.06
Years Post Secondary Teaching	14			13	15	2.4*
School Population	22337			2503	32166	37.85***
Average Class Size Taught	35.2			22	42	6.92***

* $p < .05$; ** $p < .01$; *** $p < .001$; otherwise, not significant at the .05 level

Note: Information retrieved from: [†]IPEDS; ^{††} Faculty Demographics and US Census

Research Question 1

What are the faculty perceptions of uses of Student Evaluations of Teaching?

There was one question in the survey which asked what faculty perceived to be the purpose of SET in their institution. Approximately two thirds of participants indicated that the purpose of the SET in their institution was not only for improving instruction, but also for use in making high-stakes decisions such as annual evaluations, and promotion and tenure assessments (see Table 2). There was a significant difference between ICUF and SUS faculty's reports of the purpose of SET at their institutions, as a larger percentage of the SUS faculty report SET being used for the improvement of teaching as well as FAPT decisions, as more SUS faculty report SET being used for the improvement of teaching as well as FAPT decisions. $\chi^2 (2, N = 702) = 10.07, p = .002$. A slightly more significant difference was observed between faculty when compared by tenure status, $\chi^2 (2, N = 689) = 16.22, p < .001$. In this analysis it was observed that a larger percentage of tenured faculty (68.4%) reported that SET ratings were used to improve teaching and for FAPT decisions, compared to faculty on a tenure track (61.6%) and those not on a tenure track (56.9%). The faculty who are not on a tenure track (17.7%) were more likely to report that SET ratings were used primarily for the improvement of teaching than the tenured faculty (1.3%) and the faculty on a tenure track (3.4%) Although the majority of respondents are at least somewhat in agreement with including SET ratings in making decisions about promotion and tenure, their responses indicate that they are not confident that students possess the skills and capabilities to effectively rate teaching (see Table 3, # 9).

While half of the participants do not find SET to be an effective method of evaluating professors' performance, still they were not opposed to the inclusion of SET ratings in their annual assessments. They do however, in their comments, recommend that it is only included as a part of a more comprehensive assessment. Based on faculty's responses on the survey and the comments they provided, it appears that the opposition is not to the inclusion of SET ratings in their assessments, rather to the level of impact it has on the assessment. This perceived weight of SET, which is explored in Research Question Two, is important, when two thirds of participants believe that students' ratings on the SET are motivated by the grades they anticipate receiving in the class (see Table 3, # 7).

Table 2

Uses of Student Evaluations of Teaching

	State Institution	Private Institution	Total	<i>df</i>	χ^2
What is the primary use for the Student Evaluations at your institutions?				5	52.06***
Improve Teaching	2.5%	13.5%	6.1%		
Annual Review	6.1%	11.3%	7.8%		
Promotion & Tenure Assessment	4.9%	3.5%	4.4%		
FAPT	18.6%	7.8%	15.1%		
All the above	65.5%	59.1%	63.4%		
Other	2.3%	4.8%	3.1%		

*** $p < .001$

Table 3

Perceptions of Student Evaluations of Teaching

	Strongly disagree	Disagree	Somewhat agree	Agree	Strongly agree
	%	%	%	%	%
1. Student Evaluations of Teaching ratings impact my grading of students.	38.1%	31.4%	15.8%	10.8%	3.8%
2. Other faculty consider the Student Evaluations of Teaching ratings as they assign student grades.	13.0%	26.1%	29.0%	22.9%	9.0%
3. Faculty consider the Student Evaluations of Teaching as they prepare student assignments.	11.2%	27.0%	31.6%	24.7%	5.6%
4. Some researchers indicate that professors try to get higher Student Evaluations of Teaching ratings by giving students higher grades and/or less work. I believe this actually occurs.	7.9%	22.9%	25.1%	28.6%	15.6%
5. I have considered my desired Student Evaluations of Teaching ratings when preparing student assignments.	30.9%	31.9%	19.5%	11.9%	5.9%
6. I have considered my desired Student Evaluations of Teaching ratings when grading student assignments.	42.5%	31.5%	14.2%	8.7%	3.1%
7. Students' ratings on Student Evaluations of Teaching are motivated by the grades they anticipate receiving in their class.	4.7%	16.8%	32.6%	27.5%	18.4%
8. Student Evaluations of Teaching ratings accurately reflect a professor's performance.	14.6%	32.1%	38.0%	14.2%	1.1%
9. Based on my interaction with students and my understanding of Student Evaluations of Teaching, I believe students have the skills and capabilities to effectively rate teaching.	11.3%	28.8%	35.5%	21.5%	2.9%
10. I agree with the use of Student Evaluations of Teaching as a method of effectively evaluating classroom instruction.	12.2%	23.6%	38.2%	22.3%	3.7%
11. I agree with the inclusion of Student Evaluations of Teaching ratings in annual faculty assessments.	6.5%	15.8%	39.3%	31.9%	6.6%
12. I agree with the inclusion of Student Evaluations of Teaching ratings in promotion decisions.	7.5%	17.9%	38.7%	30.5%	5.3%
13. I agree with the inclusion of Student Evaluations of Teaching ratings in tenure decisions.	8.3%	16.9%	37.9%	30.5%	6.4%

Research Question 2

How much impact do faculty believe that Student Evaluations of Teaching have on FAPT decisions?

There were four items in the survey that that were used to address this question. The details of faculty's responses to these questions are depicted in Table 4. When asked whether they have been explicitly advised that their SET ratings will impact their annual evaluation, more than half of the respondents reported that they had. Of those who responded to this follow up question, half of them also report that their SET ratings affect annual pay increases. Two thirds of the participants reported that they have been advised that SET ratings will impact their assessments for promotion and tenure decisions. Administrators from the 39 institutions invited to participate were asked what percentage of the total faculty assessment was based on SET ratings. The majority of these respondents (82%) stated that the weight of the SET ratings was not percentage-based or that it varied. When specific percentages were provided, they ranged from 25% to 50%. The same question was posed of faculty, and nearly half of respondents believed that their SET ratings account for 40% or higher on their annual evaluations, as well as assessments for promotion and tenure. Faculty's overall perceived mean weight of SET ratings on annual assessment was approximately 40%.

Table 4

Awareness of uses of SET

		State Institution	Private Institution	Total	<i>Df</i>	χ^2
Does your institution explicitly state that your Student Evaluations of Teaching ratings will impact your annual evaluation?	Yes	66.5%	49.5%	66.5%	1	12.95***
If you answered yes to the question above, do Student Evaluations of Teaching ratings have an impact on the amount of your salary increase?	Yes	53.4%	42.6%	50%	1	4.72*
Does your institution explicitly state that your Student Evaluations of Teaching ratings will impact your assessments for promotion and tenure decisions?	Yes	75.6%	58.5%	69.9%	1	21.2***

* $p < .05$, ** $p < .01$, *** $p < .001$

Research Question 3

Do faculty believe that Student Evaluations of Teaching have an impact on Grade Distribution?

Though many faculty members deny engaging in this practice, they do believe that it does occur (see Table 3, # 2-6). When asked whether SET ratings impact their own grading (Table 3, questions 1, 6), more than two thirds of participants stated it did not. When asked if they believe other faculty's grading are impacted by SET ratings, the number of respondents who disagreed were close to the number who agreed. In a telephone conversation, one professor in particular pointed out that professors are also pressured to give higher grades in order to maintain enrollment in their classes, as was reported in Voge and Higbee (2004). In elective courses especially, he stated, higher grades are important since there are no graduation requirements mandating student enrollment. Many professors who added or emailed comments suggested that, contrary to what has been reported in the literature, they do not give higher grades than deserved. Instead they adjust their teaching styles to increase interest and learning, which in turn has the potential to lead to students earning higher grades and should be expected from faculty as part of the job.

In addition to the frequencies, several tests of associations were conducted on the survey data. A significant correlation was observed between the belief that faculty in general consider SET ratings in grading and preparing assignments, and respondents actually engaging in the use of these methods. Several correlation analyses were conducted and are reported in Table 5. The purpose of these tests was to determine the

extent to which faculty's agreement that their colleagues' grading and assignment decisions are influenced by their student ratings can be associated with the respondents' partaking in that practice. The results of the correlation analysis illustrate that faculty who believe that other faculty consider their SET ratings when preparing or grading student assignments are more likely to respond that they too consider their desired SET ratings when planning and grading assignments. These correlations were significant at the .01 level.

Faculty who reported believing that other faculty members consider SET ratings in preparing student assignments were more likely to report belief that other faculty also consider the SET rating when grading, with a .51 correlation coefficient, $p < .001$. Similarly, a significant correlation of .66, $p < .001$ was found between faculty's report of having considered SET rating when preparing their own assignment and reports of considering SET rating when grading.

Several inverse relationships were observed with the participants' post secondary teaching experience. A small but significant inverse correlation was found between years of experience and faculty's report that SET ratings impact their grading, ($p < .05$). Faculty's reports of having considered SET ratings in preparing student assignments as well as grading revealed slightly stronger correlations with experience ($p < .01$). While all these correlations were significant, the correlations were very weak, with coefficients below .2. A significant correlation was not observed between years of experience and the belief that faculty in general consider SET ratings in grading. Institutional rank, tenure status, and school types did not yield any significant correlations.

Table 5

Pearson Correlation between variables

	How many years have you been teaching at a four-year college/university?	I have considered my desired Student Evaluations of Teaching ratings when preparing student assignments.	I have considered my desired Student Evaluations of Teaching ratings when grading student assignments.	Faculty consider the Student Evaluations of Teaching ratings as they assign student grades.	Faculty consider the Student Evaluations of Teaching as they prepare student assignments.
Student Evaluations of Teaching ratings impact my grading of students.	-.087*	.524***	.734***	.556***	.425***
I have considered my desired Student Evaluations of Teaching ratings when preparing student assignments.	-.104**	--	.660***	.379***	.606***
I have considered my desired Student Evaluations of Teaching ratings when grading student assignments.	-.116**	.660***	--	.521***	.476***

* $p < .05$; ** $p < .01$; *** $p < .001$; otherwise, not significant at the .05 level

Research Question 4

Are faculty interested in a unified grading system?

If so, who should be responsible for it?

Faculty who participated in this survey do not agree that a unified grading system could reduce the impact the SET has on grade distribution, as 63% of respondents indicated this. Even those who do believe that it could potentially reduce the impact that SET have on grade distribution don't agree that it should be done, as many of them added in their comments. One major concern is that a unified grading system could compromise academic freedom. Another concern is that faculty are the best qualified to grade their classes. Classes can also differ greatly from one another, which limits the applicability of a unified system.

Faculty were also asked who should be included in the development of a unified grading system if it were to be implemented. More than 83% feel that faculty should be included in the development of such system. Many participants also proposed the involvement of students in the process. In their comments, faculty stress that this would not be practical as there are too many implications. In a discussion with a professor, he expressed that one way to reduce SET influenced grading may be the implementation of more stringent practices, "but all it takes is for one professor to break rank, and the entire system could collapse". Other faculty also shared this concern in their comments, which was corroborated by findings in Research Question 3. The improvement of the SET instruments themselves and training students to be better assessors are other recommendations faculty added.

Hypothesis 1

ICUF faculty will experience greater pressure to have high SET ratings than SUS faculty.

The majority of SUS faculty reported the impact of SET ratings on their annual evaluation ranged between 20 and 50%. The ICUF faculty response, however, were more evenly spread across the continuum, reporting between 10% and 80% to be the weight of the SET ratings on their annual assessments. The perception of the weight of the SET on promotion and tenure decisions was again more widely spread among the majority of ICUF faculty, who reported between 20 and 50%, compared to SUS faculty who reported between 20 and 40%. The mode response for the two groups was the same on both questions. The difference between the ICUF and SUS faculty's responses were found to be statistically significant in a *t*-test (see Table 7). The responses across the tenure ranking did spread across the continuum as well; however, the chi-square analyses comparing these groups were also not statistically significant.

To continue with the tests, a Chi-square analysis was conducted to explore differences between ICUF and SUS faculty's reports of having been advised that SET ratings would impact their annual assessments, as well as assessments for promotion and tenure decisions. A significant difference was observed between the two groups' on this. The comparison between ICUF and SUS faculty who have been advised that SET rating impact their pay increases showed significant difference as well. Finally, there was also a statistically significant difference between ICUF and SUS faculty's reports of having been advised that SET ratings impact their assessments for promotion and tenure decisions.

The mean of the perceived weight of SET on annual assessment as well as on promotion and tenure decisions was higher for ICUF faculty, as is depicted in Table 7. The inverse relationship supports that the higher perceived weight of SET reported by ICUF faculty was significant in comparison to their SUS counterparts.

The results of the analyses reported above show that there are statistically significant differences between the variables tested, including school type, which was used as a grouping variable. In Table 4, we observe that a larger percentage of SUS faculty report that they have been advised that SET ratings impact their annual assessment and promotion and tenure decisions. Responses reported in Table 7 show that nearly two thirds of ICUF faculty perceive that SET ratings account for 40% or higher on their annual assessment, while this report is made by fewer than half of SUS faculty. This finding does support the hypothesis that ICUF faculty experience greater pressure to have higher SET ratings than their SUS counterparts. In Table 7, we also observe that a smaller percentage of ICUF faculty report that SET ratings account for 40% or higher in their assessments for promotion and tenure compared to the SUS faculty. This finding could challenge the support of the hypothesis. However, since promotion and tenure were combined in one question, and 51% of ICUF faculty report that their institution does not have a tenure system, compared to fewer than 6% of SUS faculty, the null hypothesis is rejected.

Table 7

Perceived Weight of SET Rating

	<i>t</i>	School Type	N	Mean	Std. Deviation	Std. Error Mean
How much weight do you believe is given to the Student Evaluations of Teaching ratings in your annual evaluations?	-4.708***	State Institution	465	37%	21.93	10.2
		Private Institution	226	45%	22.53	15
How much weight do you believe is given to the Student Evaluations of Teaching ratings in promotion and tenure decisions?	-5.032***	State Institution	461	37%	21.28	9.90
		Private Institution	221	46%	23.29	15.7

***p < .001

		Is your current institution a		
		State Institution	Private Institution	Total
How much weight do you believe is given to the Student Evaluations of Teaching ratings in your annual evaluations?	10%	13.3%	9.7%	12.2%
	20%	22.4%	10.6%	18.5%
	30%	23.4%	16.8%	21.3%
	40%	10.8%	14.2%	11.9%
	50%	10.3%	16.4%	12.3%
	60%	4.7%	10.6%	6.7%
	70%	7.1%	8.0%	7.4%
	80%	4.9%	10.2%	6.7%
	90%	2.2%	3.1%	2.5%
	100%	.9%	.4%	.7%
How much weight do you believe is given to the Student Evaluations of Teaching ratings in promotion and tenure decisions?	10%	12.6%	8.1%	11.1%
	20%	18.0%	12.7%	16.3%
	30%	28.9%	17.6%	25.2%
	40%	11.7%	12.2%	11.9%
	50%	10.6%	14.5%	11.9%
	60%	5.2%	9.5%	6.6%
	70%	4.8%	8.1%	5.9%
	80%	4.6%	13.6%	7.5%
	90%	3.0%	3.2%	3.1%
	100%	.7%	.5%	.6%

Hypothesis 2

SUS faculty will report less impact of SET on their grading practices than their ICUF counterparts.

Based on the ICUF (2004) report indicating the great differences between the sizes of ICUF and SUS student populations, it was anticipated that ICUF faculty ($M = 21.58$, $SD = 11.35$) would report significantly smaller class sizes than SUS faculty ($M = 61.26$, $SD = 381.49$), $t(472) = 2.56$, $p < .03$ (two tailed), $d = 39.65$. Faculty's perception of the validity of the SET was explored in Research Question 1. To begin testing this hypothesis, a Chi-square test for independence was conducted with school type as the grouping variable, and the test returned no significant associations. When faculty were asked about having considered SET ratings in preparing student assignments and grading, the majority denied engaging in this practice. The Chi-square test revealed that, over all, SUS faculty did not report less SET impact on grading than ICUF faculty, and what little difference there was between the two groups was not significant. As was found in Research Question 3, the most influential factor on a faculty's report of SET impact on their own grading was the belief that other faculty's grading was influenced by SET.

Comparing by tenure status, there was a significant difference in faculty's reports of having considered SET ratings when preparing assignments, $\chi^2(2, N = 686) = 10.92$, $p = .004$. The difference between ICUF and SUS faculty's reports of considering SET ratings when grading was also not significant, but significant differences were observed between faculty of different tenure status, $\chi^2(2, N = 686) = 11.06$, $p = .004$. Finally, the participants' tenure status also returned a significant difference concerning the impact of SET ratings on their grading overall, $\chi^2(1, N = 691) = 10.72$, $p = .005$. (See Table 6).

This hypothesis was not supported, as it appears in support of the null hypothesis, it was SUS faculty who were reporting slightly more SET impact on their grading than ICUF faculty, and the difference was not significant. Though not hypothesized, faculty's tenure status emerged as the factor having a significant association with the SET impact on grading. There were no significant associations between school type or tenure status with the belief that some professors may give students higher grades or less work in an effort to secure higher SET ratings.

Table 6

Perceived SET Impact on Grading

		Which of the following best describes your tenure status?		
		Not on a tenure track	On a tenure track	Tenured
Student Evaluation of Teaching ratings impact my grading of students.	Strongly disagree	37.3%	28.7%	43.5%
	Disagree	34.4%	33.1%	28.9%
	Somewhat agree	11.3%	20.8%	15.9%
	Agree	11.8%	13.5%	8.6%
	Strongly agree	5.2%	3.9%	3.0%
I have considered my desired Student Evaluation of Teaching ratings when preparing student assignments.	Strongly disagree	30.3%	25.6%	33.1%
	Disagree	34.1%	24.4%	34.8%
	Somewhat agree	16.1%	27.8%	17.7%
	Agree	14.2%	12.5%	10.4%
	Strongly agree	5.2%	9.7%	4.0%
I have considered my desired Student Evaluation of Teaching ratings when grading student assignments.	Strongly disagree	41.7%	33.9%	47.0%
	Disagree	35.1%	31.1%	29.5%
	Somewhat agree	10.4%	18.6%	14.8%
	Agree	9.0%	11.3%	7.0%
	Strongly agree	3.8%	5.1%	1.7%

Hypothesis 3

Class size will be inversely related to perceived importance of SET ratings.

This hypothesis was tested using a bivariate correlation analysis (see Table 8). The analysis was conducted using a Pearson correlation between the average class size reported by participants and how much weight they believed SET ratings have in the annual evaluations, as well as the weight on assessments for promotion and tenure. The correlation matrix revealed that there was virtually no correlation between class size and the weight faculty believe is assigned to SET ratings in annual faculty assessments, and the correlation was not significant. The weight on assessments for promotion and tenure decisions was equally small and also not significant. Although the analysis showed that there was an inverse relationship between class size and faculty's perceived importance of SET ratings on promotion and tenure decisions, the strength of the correlation was not sufficient to support the hypothesis. The correlation was also not significant, indicating that a true relationship cannot be inferred. This hypothesis is therefore rejected since the relationship was not supported even though the direction of the relationship was.

The correlation analysis displayed in Table 8 was conducted using the size of the student population in place of class size. A significant correlation was found between the size of the student population and the weight believed to be assigned to SET ratings in assessments for promotion and tenure. The correlation between the perceived weight of SET in annual faculty assessments and the student population was slightly weaker, though statistically significant. This supported an inverse relationship between the perceived importance of SET ratings and student population. Although this relationship

was not hypothesized it does lend support to the relationship between perceived importance of SET ratings and institution population at large.

Table 8

Class size correlations

	What is the undergraduate population at the institution at which you currently teach?	What is the average number of students in the classes you have taught at this institution?
How much weight do you believe is given to the Student Evaluations of Teaching ratings in your annual evaluations?	-.219***	.092
How much weight do you believe is given to the Student Evaluations of Teaching ratings in promotion and tenure decisions?	-.225***	-.091

*p < .05; **p < .01; ***p < .001; otherwise, not significant at the .05 level

Hypothesis 4

Faculty who perceive the importance of SET on FAPT to be high are more likely to consider SET ratings in preparing student assignments and grading.

Hypothesis 4 was also tested using bivariate correlations, depicted in Table 9. The analysis was conducted to explore the relationship between the perceived weight of SET on annual assessments and consideration of SET ratings in preparing student assignments. A significant correlation was found between these two factors. The relationship between perceived importance of SET on annual assessments and the consideration of SET ratings when grading student work was also explored. A positive significant relationship was also found between these factors. The strongest relationship however was between the perceived weight of SET ratings on annual assessment and faculty's report that SET ratings impact their grading. All tests had high levels of significance.

When the Pearson correlation analysis was performed again using the perceived weight of the SET ratings on assessments for promotion and tenure decisions, the strength of the relationships were slightly weaker. The perceived weight of the SET ratings on promotion and tenure decisions showed a positive correlation with faculty's reports of considering SET ratings when preparing student assignments, and the correlation was significant. Consideration of SET ratings in grading also had a significant correlation. For faculty reporting that SET ratings did impact their grading the correlation coefficient was slightly stronger.

The correlations found through these analyses have supported the hypothesis. Again, the correlation coefficients were very low indicating weak associations. Despite

the weak correlations, the level of significance in all these correlations was high with $p < .01$. Based on the results of these analyses, the hypothesis is supported and the null hypothesis rejected.

Table 9

Pearson Correlation between perceived importance of SET and faculty behaviors

	Student Evaluations of Teaching ratings impact my grading of students.	I have considered my desired Student Evaluations of Teaching ratings when preparing student assignments.	I have considered my desired Student Evaluations of Teaching ratings when grading student assignments.
Does your institution explicitly state that your Student Evaluations of Teaching ratings will impact your annual evaluation?	.088*	.122**	.081*
If you answered yes to the question above, do Student Evaluations of Teaching ratings have an impact on the amount of your salary increase?	.053	.068	.088
How much weight do you believe is given to the Student Evaluations of Teaching ratings in your annual evaluations?	.174***	.161***	.164***
Does your institution explicitly state that your Student Evaluations of Teaching ratings will impact your assessments for promotion and tenure decisions?	.080	.075	.069
How much weight do you believe is given to the Student Evaluations of Teaching ratings in promotion and tenure decisions?	.170***	.136***	.139***
Student Evaluations of Teaching ratings impact my grading of students.	--	.524***	.734***

* $p < .05$; ** $p < .01$; *** $p < .001$; otherwise, not significant at the .05 level

Summary

Hypothesis 1, ICUF faculty will experience greater pressure to have high SET ratings than SUS faculty, was supported by the analyses performed, which revealed that the faculty at ICUF institutions were experiencing greater pressure to have higher SET ratings. With 62% of ICUF faculty compared to 41% of SUS faculty reporting the impact of SET ratings on their annual assessments to be 40% or higher, there was a significant difference between the groups. The results of the T-test supported these associations.

Hypothesis 2 was not supported, as the findings showed significantly more SUS faculty reported that SET ratings are more highly weighted in annual evaluations than ICUF faculty. The results of Chi-square analyses revealed that the difference noted between the participants based on school setting was not significant. Supplementary analyses revealed that it was the participants' tenure status which had a significant effect on the report of whether their grading was influenced by SET ratings.

Hypothesis 3 was also rejected as class size appears to be unrelated to perceived importance of SET ratings, though those working in an institution with smaller student populations did perceive significantly greater weight attributed to SET ratings in their FAPT decisions. There, a significant inverse association was observed, $p < .01$.

Though the correlations were weak, hypothesis 4 was also supported as correlation analyses showed a positive relationship between the perceived importance of SET ratings and the impact of SET ratings on faculty's grading. These associations were all significant, $p < .01$, despite the weakness in the strength of the associations.

In all the results, although the correlations were significant at the .05 and, in most cases, .01 level, the correlations were very weak. While this research was conducted with a large sample size, there is still the possibility that the correlation could change with an increase in sample, as sample size has the power to detect small effects.

CHAPTER V

DISCUSSION

Summary

This non-experimental, correlation study was designed to explore Florida faculty's attitude toward the use of student evaluations of teaching in faculty assessments, promotion and tenure decisions, and its impact on grade inflation. Administrators at Florida's 11 State Universities and 28 Independent Colleges and Universities of Florida were contacted to invite faculty from these institutions to participate. Faculty from 26 of these institutions were invited to participate, as the others either declined or did not have an email list online.

The participants in this study, 84% of whom hold a doctorate degree, were 54% males. SUS and ICUF faculty were represented at a ratio of 2:1. Instructors accounted for 8% of the participants while Assistant, Associate and full Professors, represented approximately 30% of the sample, respectively. Eighty percent of the participants were from a tenure granting institution. Half of respondents indicated that their institutions used a standardized SET form. Nearly a third did not know what form was used, and a small percentage reported that their SET form was developed by the institution.

Interpretations

Research Question 1: What are the faculty perceptions of the uses of Student Evaluations of Teaching?

Based on the literature reviewed, which pointed to the pressure faculty experienced to get high SET ratings, the assumption was that faculty would be opposed to the use of SET in their evaluations. Nearly half of the participating faculty (47%)

report that SET ratings are not an accurate method of evaluating faculty performance, as was indicated by Baldwin (2004), Becker (2000), Edwards (2000), Germain and Scandura (2005), Nagle, (1998), and Voge and Higbee (2004). Forty-seven percent of the participants also believe that SET ratings are motivated by the grades the students anticipate receiving, a belief which was widely supported in the literature (Baldwin, 2004; Ellis, Burke, Lomire, & McCormack, 2003; Greenwald & Gilmore, 1997; Grimes et al., 2004; Isely & Singh, 2005; Millea & Grimes, 2002; Nerger, Viney, & Riedel II, 1997). Birnbaum (2000) and McAlpine and Harris (2002) suggested that students may not be prepared or possess the skills required to effectively complete the SET, and 40% of the survey participants concurred.

This study revealed a preponderance of agreement with the inclusion of SET ratings in FAPT decisions even though only 26% of the respondents agree with the use of SET as a method of effectively evaluating classroom instruction (see Table 3, # 8, 11-13). Faculty appear to want the input from their students, and many of them did suggest that students should be included in the process if a unified grading system was being developed. It is unclear, however, what level of involvement they want from the students. As is depicted in Table 3, there is some inconsistency between faculty's agreement with the inclusion of SET ratings in FAPT decisions and their lack of confidence in students' ability to perform the task of evaluating teaching.

Nerger, Viney and Riedel (1997) declared "there is never a time when student evaluations should not be used" (p. 228). The ability to show areas for teaching improvements, and ability to show teaching improvements achieved over time are some of the benefits of the individualized approach the authors propose. These two potential

benefits may in some ways be the reason that so many faculty agree to the use of SET ratings in FAPT decisions, despite their lack of confidence in the evaluators' ability to effectively perform such a task.

Research Question 2: How much impact do faculty believe that Student Evaluations of Teaching have on FAPT decisions?

Sixty-three percent of participants report that SET ratings in their institution were used for improvement of teaching as well as annual reviews and assessments for promotion and tenure, while 15% reported that it was only used for the latter two. There were also 2% of respondents who indicated they were not certain how the SET were used on their campus or that SET did not have a purpose. As reported in chapter 4, nearly half of the participating faculty believe that SET ratings account for 40% or higher in their assessments, while administrators contend that they are not percentage-based.

Teaching effectiveness should be, without a doubt, one of the primary concerns of teaching faculty and administrators in institutions of higher learning, and should receive great emphasis. Appropriate assessment of effective teaching, however, should be just as important. Nerger, Viney and Riedel (1997) propose “an individualized approach to instructor evaluation” (p. 229), and other researchers (Arreola, 2000; Marsh & Roche, 1997) support the use of a multidimensional assessment method. Student grades have in the past been another method of evaluating teaching effectiveness. While Nerger, Viney and Riedel's individualized method has the potential to address some of the concerns present in the use of SET, its applicability is debatable. A benefit of SET is that the institution only has to make a single investment to evaluate teaching. The multidimensional assessment would require additional investments, but once the system is established, the

efforts to maintain it would be limited. In an individual assessment system, however, the constant changes in academic requirements and student needs mandate changes in the course. These would require constant adjustment and investment in individualized method. One of the primary theories on grade inflation is that faculty may be giving students higher grades because they believe students' ratings on SET are motivated by their grades (Birnbaum, 1977; Crumbley et. al., 2001; McSpirit et. al., 2000). Eiszler (2002) and Millea and Grimes (2002) along with many other researchers did report that SET ratings can be impacted by the students' earned or anticipated grades. Forty-six percent of the faculty who participated in this study also believe that SET ratings are motivated by students' anticipated grades. If faculty are indeed inflating grades to get the higher SET ratings, the validity of grades as a method of assessing effective teaching is also negated. In lieu of two assessment methods which can be interdependent, thus rendering both ineffective, there is a need for the implementation of more objective and independent methods. The inclusion of such methods, or redistribution of the weight assigned to each construct in the existing evaluation systems, could diminish the actual or perceived impact of SET ratings on FAPT. Such a change may also moderate the pressure that contributes to the exclusion of important coursework or grade inflation to secure high SET ratings.

Research Question 3: Do faculty believe that Student Evaluations of Teaching have an impact on Grade Distribution?

In their responses to the survey, 32% of participants believed that other faculty do consider SET ratings when preparing assignments as well as in grading. Approximately 18% of the participants reported having considered their own SET rating when planning

or grading students' assignments. Despite these reports, there was a significant correlation observed between participants' belief that other faculty engage in these behaviors and reports of their own engagements. The correlations were moderate in strength, with significance ($p < .001$). These findings suggest that the more faculty believe that others are considering SET rating when preparing or grading student assignments, the more likely they are to also engage in such behaviors.

With these trends, the concern over applicability of a unified grading system is questionable. If faculty are more likely to consider SET ratings in preparing and grading student assignments when they believe that other faculty are also doing this, the probability remains that even with a unified grading system the trend may continue.

Whether faculty have become more lenient with the amount or quality of work that is expected of college students has some serious implications for higher education and society at large. Society, including employers and consumers of services, expect students to attain a certain level of preparedness through higher education. If course rigor is reduced to make the course more pleasing to students, than the higher education system is failing to fulfill its societal obligation of educating students, as they may not be exposed to all the material they need to become familiar with. With lenient grading practices, the institutions may fail to uphold their commitment to equip students with the knowledge to enter their chosen careers. Neglecting to uphold the high standards that are expected of institutions of higher education is a disservice to society. Students are the most affected, because in addition to being deprived of their full learning opportunities they may also be recipients of services from ill-prepared providers.

Research Question 4: Are faculty interested in a unified grading system?

If so, who should be responsible for it?

A vast majority of the participants did not believe that a unified grading system could alleviate the impact that the uses of SET have on grade distribution. It was surprising that 38% of the participants did. Even more surprising was the number of participants wrote in that although they believe that the unified system could assuage the problem, they do not believe a unified system should be implemented. That 84% of respondents recommended that faculty should be included in the development of a unified grading system was also unexpected. These results did provoke some thought on the lack of faculty inclusion in the current system of assessments and evaluations. In the literature reviewed, there was a notable absence of research on the involvement of faculty in the assessment process. Faculty involvement in peer-evaluations has been suggested, although Arreola (2000) cautions that "Care must be exercised in designing peer evaluation systems to ensure that peer judgments are not influenced or confounded by irrelevant factors" (p. 65). This supports the statement that there is not a single method of evaluating teaching and reinforces Marsh and Roche's (1997) recommendation that a multidimensional system should be used since no one method can effectively capture the essence of teaching on its own.

***Hypothesis 1:** ICUF faculty will experience greater pressure to have high SET ratings than SUS faculty.*

A 2004 ICUF report indicated that the total ICUF enrollment was 11.5% less than the combined enrollment at the three largest SUS institutions. Therefore, ICUF institutions were expected to have a greater need to employ relationship management in order to facilitate consumer [student] retention. The hypothesis that ICUF faculty will

experience greater pressure to have high SET ratings than SUS faculty was based on the concept of student consumerism and consumer retention. Together the 28 ICUF institutions award close to one third of the degrees earned in Florida every year, while there are only eleven SUS institutions competing to capture the remaining two thirds. In addition to competing with SUS's ability to accommodate many more students, ICUF institutions also must compete with the 27 other members who provide comparable service and learning environments. In order to accomplish this, they then have to be more in tune with the needs and satisfaction of the students (Kolter & Armstrong, 1999), which can be expressed through the SET ratings.

Based upon the concept of student consumerism, the ICUF institutions, with fewer students were expected to have greater pressure on faculty to show the consumer satisfaction that can be expressed in SET evaluations. The findings in this study show that there was indeed a significant difference ($p < .001$) between the groups, with ICUF faculty believing that SET ratings are weighted significantly higher in FAPT decisions than SUS faculty (see Table 7). Therefore, Hypothesis 1 was supported.

Hypothesis 2: SUS faculty will report less impact of SET on their grading practices than their ICUF counterparts.

Continuing with the same concepts of student consumerism and the need for consumer [student] retention, particularly after Hypothesis 1 was supported, it was anticipated that SUS faculty would report less impact of SET on their grading practices than ICUF faculty. This rationale was based on the belief that ICUF faculty would experience greater pressure than SUS faculty to have high SET ratings, and would have a greater need to make adjustments to secure these ratings. One factor affecting this

finding was the significant difference in class sizes reported by the two groups. With SUS classes being larger, faculty have fewer opportunities to interact with students and the students don't have as many opportunities to experience the faculty's personalities, which were found to be a characteristic that can influence SET ratings (Berg & Lindseth, 2004). In these settings, where so many of the external characteristics are removed, those remaining, including the influence of anticipated grades, could possibly become more important. At that point, faculty may need to be more conscious of SET ratings in grading since they don't have the other aspects to rely on. Upon analysis of the data, SUS faculty reported slightly more SET impact on their grading than ICUF faculty, although the difference was not statistically significant. Therefore, Hypothesis 2 was not supported.

Hypothesis 3: Class size will be inversely related to perceived importance of SET ratings.

Marsh and Roche (1997) reported that "class size actually does affect Group Interactions and Individual Rapport in a manner that is accurately reflected in SETs" (p. 1190). Since class size affects the professors' ability to interact and relate to the students, which was ranked 3rd on Berg and Lindseth's (2004) list of ten most common characteristics of effective teaching reported by students, it was expected that class size would show a significant correlation with the perceived importance of SET. The findings in this study did not support this hypothesis, as a relationship between class size and SET ratings could not be established. Surprisingly, there was almost no correlation between class size and the perceived SET weight on annual assessments (.012) or on promotion and tenure decisions (-.024).

The literature addressed class size, but not institutional population. A significant inverse correlation was found between the population and perceived weight of SET on both annual assessments and assessments for promotion and tenure. This indicates that faculty in institutions with fewer students were more likely to perceive the weight of SET to be higher. This also supports hypothesis 1 above. Hypothesis 3 was also not supported.

***Hypothesis 4:** Faculty who perceive the importance of SET in FAPT to be high are more likely to consider SET ratings in preparing student assignments and grading.*

Adaptation, one of the two psychological mechanisms “responsible for the development of our cognitive structures”, according to Piaget, consists of assimilation and accommodation (Dacey & Travers, 1996, p. 39). When researchers throughout the literature review suggested the increased use of SET ratings in FAPT decision, it was projected that faculty, based on Piaget’s theory, would assimilate and make the appropriate accommodations to secure their positions. Accommodations are certainly necessary when more than 40% of participants believe that SET ratings account for 40% or higher in FAPT decisions (see Table 7). With this belief, faculty who wish to maintain their current positions or advance within the institution must find ways to secure the desirable SET ratings. We observed in Research Question 3 that the strongest bivariate correlation was found between participants’ consideration of SET ratings in preparing and grading student assignments was the belief that other faculty were also doing this, indicating a level of assimilation. The results of the analysis support this hypothesis.

Practical Implications

Perception by most faculty in the survey is that the SET ratings that they receive have great impact on FAPT decisions at their institutions. It also appears that they

perceive students to be unprepared to complete these high stakes assessments.

Approximately 48% of faculty who participated in this study believe that the SET ratings account for 40% or higher in their annual evaluations and assessments for promotion and tenure. In the literature reviewed, researchers were able to identify factors unrelated to teaching, such as instructor personality, ability to relate to students, student anticipated grades and professor's age, which influenced SET ratings (Ellis, Burke, Lomire, & McCormack, 2003; Germain & Scandura, 2005; Marsh & Roche, 1997; Nerger, Viney, & Riedel II, 1997; Rodriguez-Ortiz, 1980). Baldwin (2004), Becker (2000), Edwards (2000), Greenwald and Gillmore (1997), Nagle (1998), and Voge and Higbee (2004) all reported a significant correlation between anticipated grades and SET ratings. In this study, 46% of participants believed that SET ratings are motivated by the grades students anticipate receiving in the course. Other researchers have indicated that students may not be best prepared to complete these assessments (Birnbaum, 2000; McAlpine & Harris, 2002). This belief was shared by 41% of participants in this study.

There appears to be a difference of opinion between faculty and administration on the validity of SET as an assessment of teaching. Administration, it would seem, finds these measures to be valid, as they include them in the FAPT decisions. There is further variation between faculty and administration's opinions on the value of the SET ratings in FAPT decisions. Of the administrators who did respond to the initial contact sent to the targeted schools, 82% indicated that there was not a specific percentage assigned to the SET in FAPT decisions. Many of them did, however, point out that teaching was one of, if not the primary, constructs in these evaluations. Survey responses show that 45% of participants believe that 40% or more of their annual assessments are dependent on

SET ratings, and 47% report the same of their assessment for promotion and tenure assessments. It is imperative that administrators take notice that nearly half of faculty believe 40% or more of the assessments which impact their ability to maintain or advance in their positions are based on the ratings of students whom they don't believe possess the skills and capabilities to accurately and effectively assess the quality of teaching.

Conclusions

1. Although ICUF faculty appeared to be experiencing greater pressure to have high SET ratings, school settings did not have show significant correlations with whether or not faculty considered SET in preparing student assignments and grading. The number of years of post secondary teaching experience and tenure status were found to have the most significant correlations with participants' report of considering SET ratings.

McSpirit et al. (2000) suggest that "75% of non-tenured faculty concerned over grade and tenure, concede being influenced by student ratings when it comes time to grade student performance" (p. 24). The findings of this study suggests that faculty who are on a tenure track are more likely to report being influenced by SET rating than their counterparts who are not on a tenure track or are already tenured. Faculty who are on a tenure track reported that they somewhat agree 21%, agree 14%, and strongly agree 4% that SET ratings impact their grading, while tenured faculty report somewhat agree 16%, agree 9% and strongly agree 3%, and those not on a tenure track report somewhat agree 11%, agree 12%, and strongly agree 5%. While 72% of tenured faculty and faculty who are not on a tenure track either disagreed or strongly disagreed that SET ratings impact their grading, 61% of tenure track faculty had those responses. The trend indicated in McSpirit et al.

(2000) was upheld, but the rate impact was not, as 61% of the tenure track faculty report that their grading is not influenced by SET ratings.

2. Although 40% of faculty agree and 33% somewhat agree that students may not possess the skills to be able to evaluate teaching, as Birnbaum (2000) and McAlpine & Harris (2002) have suggested, they are not opposed to the inclusion of these evaluations in faculty assessments. The concern of faculty it appears, is not necessarily the use of SET ratings in FAPT decisions; rather it is the weight that they carry in such decisions. ICUF faculty reported that SET ratings have a greater impact on FAPT decisions than their SUS counter parts did. Faculty from both groups do share the opinion that there is much weight placed on the SET ratings in FAPT decision, they are not against including them as part of the process.

Many participant, while opposed to the implementation of a unified grading system, as Dresner (2005) proposed, would not oppose the inclusion of student participation in the process if it were implemented. When asked who they believed should be responsible for developing a unified grading system if it were to be developed, many participants wrote in the inclusion of students. Many respondents, however, made certain to clarify that they would not agree with a unified system, as it infringes on academic freedom and had many other implications. Other participants commented that faculty were the best prepared to assess their classes and that this task should be left to them. Without endorsing the unified grading system, Dresner's (2005) reasoning that it would remove the subjective parts of grading, and allow for grades to be based more on academic merit should be noted. Replacing the subjectivity in grading with more concrete objective criteria could address the issues being faced with grade inflation. The

implementation and monitoring of such a system however could also present other difficulties.

3. The significant correlations between faculty's reports of considering desired SET rating in preparing and grading assignments and the belief that other faculty also do this indicate that this is a systems problem. Addressing this issue would then require a systems approach (Zastrow & Kirst-Ashman, 2001) to change. This would necessitate the involvement of all related parties, including faculty and administration to explore ways to change the belief that other faculty are considering desired SET rating in preparing and grading student assignments.

Strengths of the study

The target population for this study was large, and the research topic itself incited a great deal of interest from faculty, leading to a 12% response rate. This should be compared to the 8% to 10% response rates that have been found for similar electronic surveys (Dommeyer & Moriarty, 2000; Roy & Berger, 2005). The concise instrument contributed to a 95% rate of completion among those who began the survey. The high level of interest yielded a valid data set of 704 participants who were fairly representative of the population.

An extensive review of the literature was conducted to identify the gaps which had not been addressed in the existing literature. This study investigated an area which has not received much attention in research, giving voice to the faculty perspective on the issue of SET uses in FAPT decisions. This study provided faculty with a means to report on their perceptions of the impact of SET on faculty behaviors such as grade inflation and reducing course content, which was stated in the literature.

The survey was conducted online anonymously, which allowed faculty to participate at a time that was convenient for them. The anonymous survey also provided a safer medium for expression, permitting faculty to respond honestly to the survey without much fear of ramifications. The sample in this study was representative of the population.

Limitations of the study

1. There were several limitations identified in this study. The first and most obvious is the limited scope. This study involved only faculty from Florida. The study was further limited to faculty from the 11 SUS and 28 ICUF institutions. At the end of the data collection period, faculty from 26 of these institutions were sent an invitation to participate. The exclusion of the other 13 institutions was the result of the school declining to participate (6), or the school did not respond to share an email list, and the list was not publicly available online (7).
2. The demographic breakdown in the study is limited as race was addressed but ethnicity was not. As a result there was no way to identify the Hispanic professors who participated in the survey. This also contributed to the overrepresentation of White participants in the study. While a number of the Hispanic participants selected white as their race, it is uncertain how many skipped the question or how many non-White Hispanic professors participated. There was a low response rate from participants of African-American and Asian descents. According to United States census information, 8% of Florida faculty are African-American, and 4.4% are Asian, yet in this survey only 4% of participants were African-American and 2.2% were Asian. The information

available to compare the sample to the population was limited; therefore, some of the information could not be assessed for sample representativeness (see Table 1).

3. There were three questions in the survey which asked participants whether they agreed with the inclusion of SET ratings in their annual faculty assessments, promotion decisions, and tenure decisions. The survey failed to explore the level of SET inclusion they would support in these high-stakes decisions. Further exploration is needed to uncover how much weight faculty find acceptable to assign to SET in their FAPT decisions.

Recommendations for Future Study

1. Further research is needed to compare faculty perceptions to existing data to determine whether these perceptions are congruent with the facts. On average, SUS faculty believe that SET rating account for 30% on FAPT decisions, while ICUF faculty believe it is 50%, and administration's report that it is not percentage based. A review of annual faculty evaluations and assessments for promotion and tenure at SUS and ICUF institutions could reveal the extent of a possible disconnect between faculty and administration's perception of the impact in these high stakes decisions. Due to the need to maintain privacy in faculty's human resources decisions, it may be difficult to gain access to the information necessary to conduct this research. The impact that faculty believed is placed on SET ratings in FAPT decisions is high, and further research is needed to explore their behaviors in response to this belief. This could be completed by comparing faculty's perceived impact of SET in FAPT decisions to grade distribution in their courses. Again, this research may be difficult to undertake due to privacy issues with releasing student grades.

2. Another faculty concern about the use of SET in FAPT decisions is that students may not be prepared to perform such a task (Birnbaum, 2000; McAlpine & Harris, 2002). The recommendation is for students to be better informed before they complete these assessments. It is stated in the literature that having better prepared student evaluators enhances the potential to contribute to the validity of SET ratings (AAUP, 2001). There are several uncertainties that should be addressed, including:

1. Who should provide these trainings?
2. What should be included in these trainings?
3. When should these trainings be implemented?
4. How much impact could these trainings have on the accuracy of SET ratings?

Further research is necessary to answer these questions to determine whether or not such efforts would be effective, and efficient enough to be worthy of the investment.

3. Many participants who completed the survey added comments, or emailed their comments on different aspects of the survey. A large number of these emails provided further information on what faculty meant when stating that they considered SET ratings in grading. Faculty were able to share some different techniques that they have used to increase classroom learning. Some reported that they did consider previous SET ratings in preparing assignments, and explained that this helped by generating interest in the topic and assignments. Research that explores these methods is needed to further address this issue. Beyond knowing that faculty consider SET ratings in preparing student assignments and/or grading, it would be important to know to what degree and how that contributes to the quality of their teaching. SET ratings should be considered in the preparation of assignments as they should be providing feed back to help improve the

course. This concern may be best addressed in qualitative studies when researchers are able to perform in-depth investigation of the context in which SET ratings are used in preparing assignments and grading, as well as the level of influence these documents have in the process. Further study may be conducted to measure the impact on student learning resulting from faculty's adjustments to curriculum or presentation of the material.

Conclusion

As a result of this study, it was discovered that faculty are experiencing some pressure to have higher SET ratings due to their impact on FAPT decisions. Some faculty have made curriculum adjustments, in part to address these concerns, and some have adjusted their grading. The focus should be more on the types of adjustments that are being made.

This research did not include an evaluation of student grades, but the numerous studies that were reviewed show that there is an increase in the grades college students receive today from what students received two or three decades earlier. The focus of research in this area needs to be on types of changes that faculty and their institutions are making in the classroom and campus-wide that were not in place two or three decades earlier. Today, information is more readily accessible to students, a factor which can be very helpful to students by allowing them to study and conduct research when they are most productive. Many institutions around the country are implementing tools to accommodate different learning styles, this can give more students access to additional supports to increase learning capacity. An important factor that should be pondered is

whether the learning environment is more nurturing, leading to an increased opportunity for more students to learn and earn the higher grades they receive.

REFERENCES

- Addy, N., & Herring, C. (1996). Grade inflation effects of administrative policies. *Issues in Accounting Education*, 11(1), 1-2. Retrieved February 7, 2005, from ProQuest Database.
- Algozinne, B., Beattie, J., Bray, M., Flowers, C., Gretes, J., Howley, L., et al. (2004). Student evaluation of college teaching: A practice in search of principles. *College Teaching*, 52(4), 134-141. Retrieved January 31 2005 from ERIC Database.
- American Association of University Professors. (2001). *Statement on teaching evaluation*. Retrieved February 1, 2005, from <http://www.aaup.org/statements/Redbook/rbeval.htm>
- Anonymous (2002). Grade inflation scrutinized at Harvard and elsewhere. *Academe*, 88(2), 12-13. Retrieved February 7, 2005, from ProQuest Database.
- Armstrong, J. S. (1998). Are student ratings of instruction useful? *American Psychologist*, 53(11), 1223-1224. Retrieved February 12, 2005, from PsychArticles Database.
- Arreola, R. A. (2000). *Developing a comprehensive faculty evaluation system: A handbook for college faculty and administrators on designing and operating a comprehensive faculty evaluation system* (2nd ed.). Bolton, MA: Anker Publishing Company.
- Baldwin, T. K. (2004). Student evaluations of instruction. *Phi Kappa Phi Journal*, 84(14), 26-27.
- Basinger, D. (1997). Fighting grade inflation: A misguided effort? *College Teaching*, 45(3), 88-91. Retrieved February 8, 2005, from WilsonWeb Database.
- Becker, W. E. (2000). Teaching Economics in the 21st century. *Journal of Economic Perspectives*, 14(1), 109-119. Retrieved February 3, 2005, from ProQuest Database.

- Berg, C. L., & Lindseth, G. (2004). Students' perceive of effective and ineffective nursing instructors. *Journal of Nursing Education*, 43(12), 565-568. Retrieved February 3, 2005, from ProQuest Database.
- Birnbaum, M. H. (2000). In *A survey of faculty options concerning student evaluations of teaching*. Retrieved September 26, 2005, from University of California, Fullerton Web Site: <http://psych.fullerton.edu/mbirnbaum/faculty3.htm>
- Birnbaum, R. (1977). Factors related to university grade inflation [Electronic version]. *Journal of Higher Education*, 48(5), 519-539.
- Boretz, E. (2004). Grade inflation and the myth of student consumerism. *College Teaching*, 52(2), 42-46. Retrieved February 13, 2005, from ProQuest Education Journals Database.
- Buck, D. (1998). Student evaluations of teaching measures the intervention, not the effect. *American Psychologist*, 53(11), 1224-1226.
- Centra, J. A. (1972). Evaluating college teaching. In D. W. Vermilye (Ed.), *The expanded campus: Current issues in higher education* (pp. 225-233). San Francisco, CA: Jossey-Bass.
- Centra, J. A., & Gaubatz, N. B. (2002). *SIR II: Will teachers receive higher student evaluations by giving higher grades and less work?* (Research report # 10). Princeton: NJ: Educational testing Services.
- Creative Research System (n.d.). *Sample Size Calculator*. Retrieved December 21, 2007, from Creative Research System Web site: <http://www.surveysystem.com/sscalc.htm>
- Crumbley, L., Henry, B. K., & Kratchman, S. H. (2001). Students' perceptions of the evaluation of college teaching. *Quality Assurance in Education*, 9(4), 197-207.
- Dacey, J. S. & Travers, J. F. (1996). *Human development across the lifespan*. Boston: McGraw-Hill.

- Dommeier, C. J. & Moriarty, E. (2000). Comparing two forms of an e-mail survey: Embedded vs. attached. *Market Research Society. Journal of the Market Research Society*, 42(1), 39-50. Retrieved April 24, 2008, from ProQuest Database.
- Dresner, J. (2005, January 3). *Towards a unified theory of grading*. Retrieved March 6, 2005, from EducationNews.org Web Site:
<http://www.educationnews.org/towards-a-uified-theory-of-grad.htm>
- Edwards, C. H. (2000). Grade inflation: The effects on educational quality and personal well being. *Education*, 120(3), 538-546. Retrieved February 15, 2005, from ProQuest Database.
- Eiszler, C. F. (2002). College students' evaluations of teaching and grade inflation. *Research in Higher Education*, 43(4). Retrieved July 16, 2007, from ProQuest Database.
- Ellis, L., Burke, D. M., Lomire, P., & McCormack, D. R. (2003). Student grades and average ratings of instructional quality: The need for adjustment. *The Journal of Educational Research*, 91(1), 35-40. Retrieved February 15, 2005, from ProQuest Database.
- Faculty Demographics- [Tables]. (2000). *Black Issues in Higher Education*, November, 17(20), 32-35. Retrieved December 24, 2007, from: Wilson Web Database.
- Falkenberg, S. (1996). *Grade inflation*. Retrieved March 6, 2005, from Eastern Kentucky University Web Site: <http://sbs.eku.edu/PSY/FALKENBE/grdinfla.htm>
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39, 175-191.

- Florida Atlantic University (FAU) Office of Institutional Effectiveness and Analysis.
(2004, March). *Analysis of the psychometric properties of the Student Perception of Teaching instrument*. Retrieved February 6, 2005, from Florida Atlantic University Web Site: <http://iea.fau.edu/inst/spot04>
- Germain, M. L., & Scandura, T. A. (2005). Grade inflation and student individual differences as systematic bias in faculty evaluations. *Journal of Instructional Psychology*, 32(1), 58-67. Retrieved October 4, 2005, from ProQuest Databases.
- Gillmore, G. M., & Greenwald, A. G. (1999). Using statistical adjustment to reduce biases in student ratings. *American Psychologist*, 54(7), 518-519. Retrieved February 9, 1997, from PsychInfo Database.
- Gray, P. J. (2002). The roots of assessment and research: Tensions, solutions, and research directions. In T. W. Banta & associates (Eds.), *Building a scholarship of assessment* (pp. 49-66). San Francisco: Jossey-Bass.
- Greenwald, A. G. (1997). Validity concerns and usefulness of student ratings of instruction. *American Psychologist*, 52(11), 1182-1186. Retrieved February 9, 1997, from PsychInfo Database.
- Greenwald, A. G., & Gillmore, G. M. (1997). Grading leniency is a removable contaminant of student ratings. *American Psychologist*, 52(11), 1209-1217. Retrieved February 9, 1997, from PsychInfo Database.
- Grimes, P. W., Millea, M. J., & Woodruff, T. W. (2004). Grades -- Who's to blame? Student evaluation of teaching and locus of control. *Journal of Economic Education*, 35(2), 129-147.
- Hanson, G. R. (December 1, 1998). *Grade inflation: Myth or reality* (Student affairs research at University of Texas - Austin). Retrieved February 1, 2005 from: the University of Texas website:
<http://www.utexas.edu/student/research/reports/inflation/inflation.html>

- Harvey, P. M. (1999 August 26). *The development of the student's evaluation of instructor assessment*. Retrieved March 5, 2005 from the Mt. Royal College Website: <http://www.mtroyal.ab.ca/mrfa/pubs/acepaper.htm>
- Haskell, R. E. (1997). Academic freedom, tenure, and student evaluation of faculty: Galloping polls in the 21st century. *Education Policy Analysis Archives*, 5(6), . Retrieved February 22, 05, from <http://epaa.asu.edu/epaa/v5n6.html>
- Hassel, H., & Lourey, J. (2005). The dea(r)th of student responsibility. *College Teaching*, 53(1), 2-13. Retrieved February 13, 2005, from ProQuest Education Journals Database.
- Independent Colleges and Universities of Florida (2004), October) *Florida's Independent colleges and Universities: Private institutions serving a public purpose ICUF as One University*. Tallahassee, FL: Author
- Isely, P., & Singh, H. (2005). Do higher grades lead to favorable student evaluations? *Journal of Economic Education*, 36(1), 29-42.
- Israel, G. D. (1992). PEOD6: Determining Sample Size. *University of Florida, Institute of Food and Agricultural Sciences*. Retrieved from the University of Florida: http://edis.ifas.ufl.edu/PD006#SECTION_5
- Johnson, V. E. (2003). *Grade inflation: A crisis in college education*. New York: Springer.
- Kanagaretnam, K, Mathieu, R & Thevaranjan, A. (2003), An economic analysis of student evaluations: Implications of universities. *Managerial Decision Economics*, 24(1), 1-13. Retrieved November 6, 2007, from ProQuest Database.
- Kohn, A. (2002, November 8). The dangerous myth of grade inflation [Electronic version]. *The Chronicle of Higher Education*, 49, 11. Retrieved February 1, 2005, from <http://www.alfiekohn.org/teaching/gi.htm>
- Kolter, P. & Armstrong, G. (1999). *Principles of Marketing* (8th ed.). Upper Saddle River, NJ: Prentice Hall

- Manfreda, K. L & Vehovar, V. (n.d.) Survey design features influencing response rates in web surveys. Retrieved December 26, 2007, from http://www.icis.dk/ICIS_papers/C2_4_3.pdf
- Marr, K. R. (2000). The role of teaching effectiveness in tenure and post-tenure decisions at small independent liberal arts colleges (Doctoral dissertation, Marquette University, 2000). *Dissertation Abstracts International, DAI-A 61/07, 25792776*. Retrieved October 2, 2005 from ProQuest Dissertations and Theses
- Marsh, H. W., & Roche, L. A. (1997). Making students' evaluations of teaching effectiveness effective: The critical issues of validity, bias and utility. *American Psychologist, 52*(11), 1187-1197. Retrieved February 9, 2005, from PsychInfo Database.
- McAlpine, L & Harris, R. Evaluating teaching effectiveness and teaching improvement: A language for institutional policies and academic development practices. *International Journal for Academic Development, 7*(2), 7-17.
- McSpirit, S., Kopacz, P., Jones, K., & Chapman, A. (2000). Faculty opinion on grade inflation: Contradictions about its cause. *College and University Journal, 75*(3), 19-25.
- Merritt, D. J. (2008). Bias, the brain, and student evaluations of teaching. *St. John's Law Review, 82*(1), 235-287. Retrieved April 23, 2008, from ProQuest Database
- Millea, M., & Grimes, P. W. (2002). Grade Expectations and Student Evaluation of Teaching. *College Student Journal, 36*(4), 582-590. Retrieved September 26, 2005, from WilsonWeb journal directory
- Murray, H. G. (1997). Does evaluation of teaching lead to improvement of teaching? *International Journal of Academic Development, 2*(1), 8-23.
- Nagle, B. (1998). A proposal for dealing with grade inflation: The relative performance index. *Journal of Education for Business, 74*(1), 40-43. Retrieved February 1, 2005, from ProQuest Database.

- Nerger, J. L., Viney, W., & Riedel II, R. G. (1997). Student ratings of teaching effectiveness: Use and misuse. *Midwest Quarterly*, 38(2), 218-233.
- Neusner, J. (1984). *How to grade your professors*. Boston: Beacon Press.
- Nummedal, S. G. (1994). How classroom assessment can improve teaching and learning. In D. F. Halpern & Associates (Eds.), *Changing college classrooms: New teaching and learning strategies for an increasingly complex world* (pp. 289-305). San Francisco: Jossey-Bass Publishers.
- Onwuegbuzie, A. J., Witcher, A. E., Collins, K. M. T., Filer, J. D., & Moore, C. W. (2007). Students' perceptions of characteristics of effective college teachers: A validity study of a teaching evaluation form using a mixed-methods analysis. *American Educational Research Journal*, 44(1), 113-160. Retrieved April 23, 2008, from ProQuest Database
- Parayitam, S., Desai, K., Phelps, L. D. (2007). The effect of teacher communication and course content on student satisfaction and effectiveness. *Academy of Educational Leadership Journal* 11 (3), 91-105. Retrieved April 23, 2008, from ProQuest Database
- Potter, W., Nyman, M. A., & Klumpp, K. S. (2001). Be careful what you wish for: analysis of grading trends at a small liberal arts college, grade inflation or progress. *College and University Journal*, 76(4), 9-14.
- Roarty, M. (10/25/04). Dumbing down the A's. *GW Hatchet*. Retrieved February 1, 2005 from http://www.gwhatchet.com/global_user_elements/printpage.cfm?=&779324
- Rodriguez-Ortiz, A. R. (1980). The relationship among selected environmental factors; personal characteristics of instructors, and student evaluation of instruction in an institution of higher education (Doctoral dissertation, New York University, 1980). *Dissertation Abstracts International, DAI-A 41/12*, 4999. Abstract retrieved October 2, 2005 from ProQuest Dissertations and Theses.

- Roy, A. & Berger, P. D. (2005). E-mail and mixed mode database surveys revisited: Exploratory analyses of factors affecting response rates. *Journal of Database Marketing & Customer Strategy Management*, 12(2), 153-171. Retrieved April 24, 2008, from ProQuest Database.
- Scanlan, J. M., & Care, W. D. (2004). Grade inflation: Should we be concerned. *Journal of Nursing Education*, 43(10), 475-478.
- United States Department of Commerce, Bureau of the Census. (2000) Census Sample Data File.
- University of Southern Indiana. (2002.). Student evaluation of teaching task force position paper. In *Student evaluation of teaching*. Retrieved from the University of Southern Indiana: <http://www.usi.edu/distance/set/SETpaper/rtf>
- Voge, D. J., & Higbee, J. L. (2004). A grade "A" controversy: A dialogue on grading policies and related issues. *Research and Teaching in Developmental Education*, 21(1), 63-77. Retrieved January 31, 2005, from ProQuest Database.
- Western Association of Schools and Colleges (WASC). (2001). Handbook of Accreditation. Retrieved October 20, 2005 from: http://www.aacu.org/irvinediveval/pdfs/WASC_Handbook.pdf
- Wilson, B. P. (1999). The phenomenon of grade inflation in higher education. *The phi Kappa Phi Journal*, 79(4), 38-41. Retrieved February 8, 2005, from ProQuest Database.
- Wilson, M. L. (1998). An examination of instructional effectiveness in higher education using multiple outcome measures (Doctoral dissertation, Middle Tennessee State University, 1998). *Dissertation Abstracts International, DAI-A 59/04*, 1269. Retrieved October 2, 2005 from ProQuest Dissertations and Theses
- Yunker, J. A., & Marlin, Jr., J. W. (1984). Performance evaluation of college and university faculty: An economic perspective. *Educational Administration Quarterly*, 20(1), 9-37. Retrieved January 31, 2005, from ProQuest Database.

Yunker, P. J., & Yunker, J. A. (2003). Are student evaluations of teaching valid? Evidence from an analytical business core course. *Journal of Education for Business*, 78(6), 313-317. Retrieved February 1, 2005, from ProQuest Database.

Zastrow, C., & Kirst-Ashman, K. K. (2001). *Understanding human behavior and the social environment* (5th ed.). Belmont, CA: Brooks/Cole.

BIBLIOGRAPHY

- Adler, E. S. & Clark, R. (1999). *How it's done: An invitation to social research*. Boston: Wadsworth Publishing Company.
- American Psychological Association (2001). *Publication manual of the American Psychological Association* (5th ed.). Washington, DC: Author.
- Canto, V. M.; Darr, W. & Campbell, C. A. (2007). Performance appraisal of behavior-based competencies: A reliable and valid procedure. *Personnel Psychology*, 60(1), 201-230. Retrieved December 19, 2007, from ProQuest Databases.
- Dowell, D. A., & Neal, J. A. (1982). A selective review of the validity of student ratings of teaching. *The Journal of Higher Education*, 53(1), 51-62. Retrieved September 29, 2005, from JSTOR Databases.
- Dyer, G. J. (1982). Determinants of academic promotion in undergraduate science departments (Doctoral dissertation, Rutgers The State University of New Jersey - New Brunswick, 1982). *Dissertation Abstracts International, DAI-A 43/05*, 1490. Abstract retrieved October 2, 2005 from ProQuest Dissertations and Theses.
- Kauhanen, A. & Piekkola, H. (2006). What Makes Performance-Related Pay Schemes Work? Finnish Evidence. *Journal of Management and Governance*, 10(2), 149-177. Retrieved December 19, 2007, from ProQuest Databases.
- Kolodny, A. (1998). *Failing the future: A dean looks at higher education in the twenty-first century*. Durham, NC: Duke University Press.
- Levine, D. M.; Krehbiel, T. C. & Berenson, M.L. (2000). *Business statistics: A first course* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.

- Mesmer, M. (2004). Developing effective performance reviews. *Strategic Finance*, 85(9), 13-14. Retrieved October 12, 2006, from ABI/INFORM Global.
- Mood, A. M. (1973). *The future of Higher Education: Some Speculations and suggestions*. New York: McGraw-Hill.
- Rubin, A. & Babbie, E. (2000). *Research methods for Social Work* (4th ed.). Belmont, CA: Brooks/Cole
- Sewell, E. (2004). Grade dropping: An Empirical analysis. *Journal of Economic Education*, 25(1), 24-34.

APPENDIX A
IRB Approval Letter



Lynn University

Principal Investigator: Judi Cineas

Project Title: A Study of Faculty Attitude Toward The Use of Student Evaluations of Teaching in Faculty Assessments, Promotion and Tenure decisions and its Impact on Grade Inflation

IRB Project Number: 2008-006 Request for IRB Exemption of Application and Research Protocol for a New Project

IRB Action by the IRB Chair or Another Member or Members Designed by the Chair:

Expedited Review of Application and Research Protocol and Request for Expedited Review (FORM 3): Approved ; Approved w/provision(s)

COMMENTS:

Consent Required: No Yes Not Applicable Written Signed

Consent forms must bear the research protocol expiration date of 3/12/09.

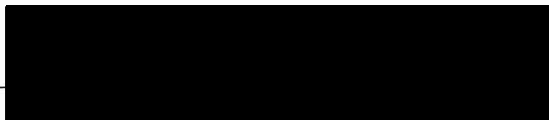
Application to Continue/Renew is due:

1) For an Expedited IRB Review, one month prior to the due date for renewal

2) Other:

Name of IRB Chair: Farideh Farazmand

Signature of IRB Chair



Date: 3/12/08

Cc. Dr. Cipolla

Institutional Review Board for the Protection of Human Subjects

Lynn University

3601 N. Military Trail Boca Raton, Florida 33431

APPENDIX B

Introductory Email to Office of Academic Affairs/Provost

Greetings Dr. :

Thank you for taking time to assist me in this research. As I mentioned on the phone, I am a student at Lynn University. I am in the process of gathering some preliminary information for my dissertation research. My dissertation topic is *A study of Faculty Attitude Toward The Use of Student Evaluations of Teaching in Faculty Assessment, Promotion and Tenure Decisions and its impact on Grade Inflation.*

I am currently interested in the following;

1. Does your school have a written policy advising instructors that the student evaluations they receive will be used in annual faculty assessments?
2. The percentage of the total Assessment that are based on student evaluations.
3. Later in the research process I may be looking at general grading (by school, not by student or by instructor) in order to explore any correlation between grades and student evaluations. I will be doing this in order to compare schools which have this written policy and those where there is not a policy.
 - a. Would your school be willing to share this information?
 - b. If so who would be the appropriate person to contact?
4. I will also be assessing faculty perceptions of the use of student evaluations of teaching. Would your institution allow me to send such a survey to your faculty?

If yes who would be the appropriate individual to contact to get a faculty email distribution list?

If I need permission from someone other than you to request the distribution list and/or send the email please provide name and contact information.

Many thanks for your time and attention. I truly appreciate your assistance.

Sincerely,

Judi Cinéas, MSW

Doctoral Candidate

████████████████████

APPENDIX C

Email Request to Office of Academic Affairs/Provosts

Greetings Dr. :

Thank you for agreeing to assist me in this research. As I mentioned in my introductory email, I am a student at Lynn University, and am in the process of collecting data for my dissertation research. My dissertation topic is *A study of Faculty Attitude Toward The Use of Student Evaluations of Teaching in Faculty Assessments, Promotion and Tenure Decisions and its Impact on Grade Inflation.*

I am hoping to invite your faculty to participate in an anonymous online survey being conducted for the purpose of this dissertation research. Per our previous email communication, I am requesting a faculty email distribution which will be used to send your faculty email invitations to participate. The survey addresses faculty's perception on the use of Student Evaluations of teaching in the annual assessment as well as evaluations for promotion and tenure decisions. The survey will be administered anonymously and responses will not be tied to any individual institution or faculty member once submitted.

Many thanks for your time and attention. I truly appreciate your assistance.

Sincerely,

Judi Cinéas, MSW

Doctoral Candidate

██████████

████████████████████

APPENDIX D

Invitation to Participate Sent to Professors

Greetings Professor:

My name is Judi Cinéas. I am a doctoral student at Lynn University. I am currently in the process of conducting the research for my dissertation, and am inviting you to participate in this project.

The topic of my dissertation is: *A study of Faculty Attitude Toward The Use of Student Evaluations of Teaching in Faculty Assessments, Promotion and Tenure Decisions and its Impact on Grade Inflation*

As you may well know there have been concerns in academia over grade inflation. Some researchers have identified the use of student evaluations in the faculty assessments as a contributing factor to grade inflation. The purpose of my research is to gather the faculty perspective. The study aims to see if faculty believe this is true and if they engage in such practices. Instructors, Assistant, Associate and full Professors at the 11 State Universities and 28 members Independent Colleges and Universities of Florida have been invited to participate in this research.

Please follow this [link](#) and type the password “**set**” to go to the consent form. By clicking the “I agree” button at the bottom of the page you will be consenting to participation in this anonymous survey on SurveyMonkey.com. I would greatly appreciate your taking 5-10 minutes from your already busy schedule to complete this 36 question survey. Your participation is voluntary and your responses are anonymous. No one will be able to identify you or your school from your survey responses. Please note that the anonymous format of this survey limits my ability to honor requests to revoke consent as I will not be able to match responses with individual participants.

The survey will only be available online until Friday April 11, 2008. The information gathered is designated for use solely in this study. I will keep only the responses to the questionnaire which SurveyMonkey transmits to a spreadsheet for 5 years. This survey has been set up so that SurveyMonkey will NOT save your IP address.

If you have any questions, please feel free to contact me, Judi Cinéas, at [REDACTED] or [REDACTED]. Thank you for your attention and time.

Respectfully,

Judi Cinéas, MSW
Doctoral Candidate
Lynn University

APPENDIX E
Follow up Email Sent to Professors

Greetings Professor:

My name is Judi Cinéas. I am a doctoral student at Lynn University, in the process of conducting the research for my dissertation. The topic of my dissertation is: *A study of Faculty Attitude Toward The Use of Student Evaluations of Teaching in Faculty Assessments, Promotion and Tenure Decisions and its Impact on Grade Inflation.*

In March, I invited faculty to participate in an anonymous online survey being conducted for my dissertation research. The purpose of my research is to gather the faculty perceptions on the use of Student Evaluations of Teaching in Faculty Assessment for merit decisions. Instructors, Assistant, Associate and full Professors at the 11 State Universities and 28 members of the Independent Colleges and Universities of Florida have been invited to participate in this research.

I thank you if you have already completed the survey. If not, I would greatly appreciate your decision to participate in the survey. Please follow this [link](#) and type the password **set** to go to the consent form leading to the anonymous survey on SurveyMonkey.com. The survey contains 36 questions and should only take a few minutes. Your participation is voluntary and your responses are anonymous. I will not be able to identify you from your survey responses. Please note that the anonymous format of this survey limits my ability to honor requests to revoke consent as I will **not** be able to match responses with individual participants.

The survey will only be available online until **Friday April 11, 2008**. The information gathered is designated for use in this study. I will keep only the responses to the questionnaire which SurveyMonkey transmits to a spreadsheet for 5 years. Survey monkey will NOT save your IP address. All other information will be discarded upon completion of the study.

If you have any questions, please feel free to contact me, Judi Cinéas at [REDACTED] Again, thank you for your attention and time.

Respectfully,

Judi Cinéas, MSW
Doctoral Candidate
Lynn University

APPENDIX F

Electronic Survey Instrument (including Consent form)

Faculty Perception of Uses of Student Evaluations

1. Consent

Informed Consent Form

Institutional Review Board for the Protection of Human Subjects
Lynn University
3601 North Military Trail, Boca Raton, Florida 33431

Lynn University
THIS DOCUMENT SHALL ONLY BE USED TO PROVIDE AUTHORIZATION FOR VOLUNTARY CONSENT

PROJECT TITLE: A study of Faculty Attitude Toward The Use of Student Evaluations of Teaching in Faculty Assessment, Promotion and Tenure Decisions and its Impact on Grade Inflation.

Project IRB Number: 2008-006 Lynn University 3601 N. Military Trail Boca Raton, Florida 33431

I Judi Cinéas, am a doctoral student at Lynn University. I am studying Global Leadership, with a specialization in Corporate and Organizational Management. One of my degree requirements is to conduct a research study.

DIRECTIONS FOR THE PARTICIPANT:

You are being asked to participate in my research study. Please read this carefully. This form provides you with information about the study. The Principal Investigator (Judi Cinéas) will answer all of your questions. Ask questions about anything you don't understand before deciding whether or not to participate. You are free to ask questions at any time before, during, or after your participation in this study. Your participation is entirely voluntary and you can refuse to participate without penalty or loss of benefits to which you are otherwise entitled.

You acknowledge that you are at least 18 years of age, and that you do not have medical problems or language or educational barriers that precludes understanding of explanations contained in this authorization for voluntary consent.

PURPOSE OF THIS RESEARCH STUDY: The purpose of this study is to explore faculty perceptions of the use of Student Evaluations of Teaching (SET) in Faculty Assessment, Promotion and Tenure (FAPT) decisions, and determine if a relationship does exist between *Student Evaluations of Teaching* and grade inflation. More than 5000 (IRB--a definite number will be added when the email list is compiled. Most of the schools don't have this information on their websites that I could find) full time professors have been invited to participate in this study. The individuals invited to participate are all faculty members at one of Florida's 11 State Universities or one of the 28 member schools of the Independent Colleges and Universities of Florida.

PROCEDURES:

If you agree to participate after reading this consent form, please click the "I agree" button below and you will be directed to the survey. You will be able to complete the survey containing four sections (*Demographics, Institution Information, Student Evaluations and Grading, and Recommendations*), with a total of 36 questions. The survey should take 5-10 minutes to complete. SurveyMonkey is HACKER SAFE CERTIFIED. The site undergoes the HACKER SAFE security scan daily to endure privacy.

POSSIBLE RISKS OR DISCOMFORT: This study involves minimal risk. You may find that some of the questions are sensitive in nature. In addition, participation in this study requires a minimal amount of your time and effort.

POSSIBLE BENEFITS: There may be no direct benefit to you in participating in this research, but knowledge may be gained which may help contribute to the literature available on the impact that SET have on grading and grade inflation.

FINANCIAL CONSIDERATIONS: There is no financial compensation for your participation in this research. There are no costs to you as a result of your participation in this study.

Faculty Perception of Uses of Student Evaluations

ANONYMITY

This Survey will be anonymous. No names, social numbers, institutional affiliations, ID numbers, or other identifiers will be requested. The IP address of participants will NOT be saved. Invitations to participate in this survey were sent in a blind copy email so that recipients could not be identified by others. The email list will not be matched with responses. Anonymity will be maintained to the degree permitted by the technology used. Specifically, no guarantees can be made regarding the interception of data sent via the Internet by any third parties. The researcher will not identify you and data will be reported as "group" responses. Participation in this survey is voluntary. Clicking I agree to begin the survey constitutes your informed consent to participate. All information will be held in strict confidence and will not be disclosed unless required by law or regulation.

The results of this study may be published in a dissertation, scientific journals or presented at professional meetings. In addition, your individual privacy will be maintained in all publications or presentations resulting from this study.

All the data gathered during this study, which were previously described, will be kept strictly confidential by the researcher. Data will be stored as files on password protected computers until destroyed after 5 years. All information will be held in strict confidence and will not be disclosed unless required by law or regulation.

RIGHT TO WITHDRAW: You are free to choose whether or not to participate in this study. There will be no penalty or loss of benefits to which you are otherwise entitled if you choose not to participate.

CONTACTS FOR QUESTIONS/ACCESS TO CONSENT FORM: Any further questions you have about this study or your participation in it, either now or any time in the future, will be answered by Judi Cinéas (Principal Investigator) who may be reached at: [REDACTED] or via email at [REDACTED] and Dr. John Cipolla, faculty advisor who may be reached at: [REDACTED]. For any questions regarding your rights as a research subject, you may call Dr. Farideh Farazmand, Chair of the Lynn University Institutional Review Board for the Protection of Human Subjects, at [REDACTED]. If any problems arise as a result of your participation in this study, please call the Principal Investigator (Judi Cinéas) and the faculty advisor (Dr. John Cipolla) immediately.

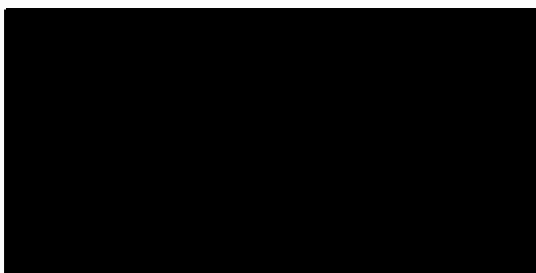
AFFIDAVIT:

I hereby certify that a written explanation of the nature of the above project has been provided to the person participating in this project. A copy of the written documentation provided is attached hereto. By the person's consent to voluntary participate in this study, the person has represented that he/she is at least 18 years of age, and that he/she does not have a medical problem or language or educational barrier that precludes his/her understanding of my explanation. Therefore, I hereby certify that to the best of my knowledge the person participating in this project understands clearly the nature, demands, benefits, and risks involved in his/her participation.

_____ **Date of IRB Approval: 2008-006 Expiration**

Date: 3/12/09

Signature of Investigator



Faculty Perception of Uses of Student Evaluations

*** 1. After having read the consent letter please click the "I agree" button if you wish to continue to the survey. If you do not wish to continue, you may exit by clicking on "exit this survey" in the upper right corner of the page.**

I agree

Faculty Perception of Uses of Student Evaluations

2. Demographics

This section will help me get some information about the participants of this research. Please respond to each question by clicking on the button preceding your response. Where a box is provided please type your numerical response.

2. What is your gender?

- 1. Male
- 2. Female

3. What is your age?

4. What is the highest educational degree you hold?

- 1. Bachelors
- 2. Master
- 3. PhD/ EdD

5. What is your annual income from teaching?

- 1. Less than \$30,000
- 2. \$30,000 -\$50,000
- 3. \$50,000 -\$75,000
- 4. More than \$75,000

6. What is your current institutional rank?

- 1. Instructor
- 2. Assistant Professor
- 3. Associate Professor
- 4. Professor
- 5. Other (please specify)

7. What racial background do you identify with?

- 1. American Indian or Alaska Native
- 2. Asian
- 3. Black or African American
- 4. Native Hawaiian or Other Pacific Islander
- 5. White

Faculty Perception of Uses of Student Evaluations

3. School Information

This section will allow me to learn about the institution where you are teaching. Please respond to each question by clicking on the button preceding your response. Where a box is provided please type your numerical response.

8. Is your current institution a

- 1. State Institution
- 2. Private Institution

9. Is your institution located in a

- 1. Rural setting
- 2. Urban setting

10. What is the undergraduate population at the institution at which you currently teach (provide best estimate)?

11. What is the average number of students in the classes you have taught at this institution?

12. How many years have you been teaching at a four-year college/university?

13. Does your institution have a tenure system?

- 1. No
- 2. Yes

14. Which of the following best describes your tenure status?

- 1. Not on a tenure track
- 2. On a tenure track
- 3. Tenured

15. Does this institution use a form of Student Evaluation of Teaching?

- 1. No
- 2. Yes

16. Which Student Evaluation form is used?

- 1. SIR II (Student Instructional Reports II)
- 2. SPOT (Student Perception of Teaching)
- 3. Other Form (please specify)

Faculty Perception of Uses of Student Evaluations

17. What is the primary use for the Student Evaluations at your institutions?

- 1. Improvement of teaching
- 2. Yearly faculty review
- 3. Faculty Assessment for Promotion and Tenure
- 4. All the above
- 5. Other (please specify)

18. Does your institution explicitly state that your Student Evaluation of Teaching ratings will impact your annual evaluation?

- 1. No
- 2. Yes

19. If you answered yes to question #18 above, do Student Evaluation of Teaching ratings have an impact on the amount of your salary increase?

- 1. No
- 2. Yes

20. How much weight do you believe is given to the Student Evaluation of Teaching ratings in your annual evaluations? Please select the answer that best applies.

- 10%
- 20%
- 30%
- 40%
- 50%
- 60%
- 70%
- 80%
- 90%
- 100%

21. Does your institution explicitly state that your Student Evaluation of Teaching ratings will impact your assessments for promotion and tenure decisions?

- 1. No
- 2. Yes

22. How much weight do you believe is given to the Student Evaluation of Teaching ratings in promotion and tenure decisions? Please circle the answer that best applies.

- 10%
- 20%
- 30%
- 40%
- 50%
- 60%
- 70%
- 80%
- 90%
- 100%

Faculty Perception of Uses of Student Evaluations

4. Student Evaluations of Teaching and Grading

In this section, I would like to discover your thoughts on the uses of Student Evaluations of Teaching. Please rate your agreement with each statement by clicking on the appropriate button.

23. Student Evaluation of Teaching ratings impact my grading of students.

- 1. Strongly disagree
- 2. Disagree
- 3. Neutral
- 4. Agree
- 5. Strongly agree

24. Other faculty consider the Student Evaluation of Teaching ratings as they assign student grades.

- 1. Strongly disagree
- 2. Disagree
- 3. Neutral
- 4. Agree
- 5. Strongly agree

25. Faculty consider the Student Evaluations of Teaching as they prepare student assignments.

- 1. Strongly disagree
- 2. Disagree
- 3. Neutral
- 4. Agree
- 5. Strongly agree

26. Some researchers indicate that professors try to get higher Student Evaluation of Teaching ratings by giving students higher grades and/or less work. I believe this actually occurs.

- 1. Strongly disagree
- 2. Disagree
- 3. Neutral
- 4. Agree
- 5. Strongly agree

27. I have considered my desired Student Evaluation of Teaching ratings when preparing student assignments.

- 1. Strongly disagree
- 2. Disagree
- 3. Neutral
- 4. Agree
- 5. Strongly agree

Faculty Perception of Uses of Student Evaluations

28. I have considered my desired Student Evaluation of Teaching ratings when grading student assignments.

- 1. Strongly disagree
- 2. Disagree
- 3. Neutral
- 4. Agree
- 5. Strongly agree

29. Students' ratings on Student Evaluations of Teaching are motivated by the grades they anticipate receiving in their class.

- 1. Strongly disagree
- 2. Disagree
- 3. Neutral
- 4. Agree
- 5. Strongly agree

30. Student Evaluation of Teaching ratings accurately reflect a professor's performance.

- 1. Strongly disagree
- 2. Disagree
- 3. Neutral
- 4. Agree
- 5. Strongly agree

31. Based on my interaction with students and my understanding of Student Evaluation of Teaching, I believe students have the skills and capabilities to effectively rate teaching.

- 1. Strongly disagree
- 2. Disagree
- 3. Neutral
- 4. Agree
- 5. Strongly agree

32. I agree with the use of Student Evaluations of Teaching as a method of effectively evaluating classroom instruction.

- 1. Strongly disagree
- 2. Disagree
- 3. Neutral
- 4. Agree
- 5. Strongly agree

Faculty Perception of Uses of Student Evaluations

33. I agree with the inclusion of Student Evaluations of Teaching ratings in annual faculty assessments.

- 1. Strongly disagree
- 2. Disagree
- 3. Neutral
- 4. Agree
- 5. Strongly agree

34. I agree with the inclusion of Student Evaluations of Teaching ratings in promotion decisions.

- 1. Strongly disagree
- 2. Disagree
- 3. Neutral
- 4. Agree
- 5. Strongly agree

35. I agree with the inclusion of Student Evaluations of Teaching ratings in tenure decisions.

- 1. Strongly disagree
- 2. Disagree
- 3. Neutral
- 4. Agree
- 5. Strongly agree

Faculty Perception of Uses of Student Evaluations

5. Recommendations

Some researchers have recommended a unified grading policy (such as grade rationing) which would provide standards for grading across the board. The intent would be to make grading less subjective, and reduce potential for grade inflation.

Please check the responses that most closely match your opinion of this recommendation.

36. Do you believe a unified grading policy could reduce the impact the Student Evaluations of Teaching has on grade distribution?

- 1. No
- 2. Yes

37. If yes, who should be responsible for developing such a system? Please check all that apply.

- 1. Faculty
- 2. Administration
- 3. Accrediting bodies
- Other (please specify)

445B14 T 4491
07/07/08 39800

Group