

To the Graduate Council:

I am submitting herewith a thesis written by Kristy Ann Carter Ouallal-McRiffey entitled "How Do You Rate Those Instructors? The Validity of Student Evaluations of Higher Education Instructors." I have examined the final electronic copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Instructional Technology and Educational Studies.

John Ray
Major Professor

We have read this thesis
and recommend its acceptance:

Gary Skolits

Edward Counts

Accepted for the Council:

Anne Mayhew
Vice Chancellor and
Dean of Graduate Studies

(Original signatures are on file with official student records.)

How Do You Rate Those Instructors?
The Validity of Student Evaluations of Higher Education Instructors

A Thesis
Presented for the
Master of Science
Degree
The University of Tennessee, Knoxville

Kristy Ann Carter Ouallal-McRiffey
December 2005

DEDICATION

This thesis is dedicated to my children Jeffery and Nadia. They keep me going when I feel like giving up. I also dedicate this thesis to my husband Albert, who is always home at the end of the day when I come home tired and feeling beat up. His support has been my rock upon which I rest.

ACKNOWLEDGEMENT

I wish to thank all those who made it possible for me to complete my Master of Science degree in Curriculum, Research, and Evaluation. Thank you, Dr. John Ray for your guidance throughout the research and writing process. Thank you also to Dr. Gary Skolits and Dr. Edward Counts for serving on my committee. I truly appreciate all your time and effort.

A special thanks to Dr. Cheryl Buehler for giving me the gift of knowledge and experience working on her research project. It has made a big difference in my life. Also thanks to Dr. Karen Franck for sharing her knowledge and statistical expertise with me. Thanks also to Dr. Julia Malia and Dr. James Malia for their support and kindness. You are all very special people. I would also like to thank each one of the participants in this study.

I would also like to thank others who have been vital players in my life that contributed to my receiving this degree. First, I would like to thank my grandmother, Lena Hutchison, for passing on her determination. I know I wouldn't be here today without it. I would also like to thank my aunt, Vita Hutchison. She was always on my side no matter what and taught me to hang in there when the going gets tough.

ABSTRACT

The purpose of this study was to examine the validity of students' evaluation of higher education instructors. Many questions and issues arise concerning students' evaluation of higher education instructors. It is questionable whether a student can provide an unbiased evaluation of a higher education instructor. Because there are varying opinions on what constitutes a good or even an effective instructor, it makes it hard to determine just what student evaluations measure. If factors such as the personal feelings toward the instructor are used to determine the score on an evaluation, then this can't possibly measure the instructors' ability to teach. Instructors who participated felt that students are not likely to provide unbiased evaluations of higher education instructors. More than half of the students in this study said that they use their personal feelings toward an instructor and the perceived grade in the course to base their opinions when evaluating an instructor. Furthermore, students also reported using the personal characteristics of an instructor to determine responses on evaluations.

TABLE OF CONTENTS

CHAPTER	PAGE
I. INTRODUCTION.....	1
Statement of the Problem.....	1
Purpose of the Study.....	2
Research Objectives.....	2
Research Questions.....	2
Assumptions.....	3
Limitations.....	3
Definition of Terms.....	3
II. REVIEW OF LITERATURE.....	4
Summary of Literature.....	10
III. METHODS.....	11
Participants.....	11
Implementation.....	12
Data Analysis.....	12
IV. RESULTS.....	13
Instructor Interviews.....	13
Student Questionnaires.....	15
V. SUMMARY, DISCUSSION, CONCLUSION, AND RECOMMENDATIONS.....	21
Summary.....	21
Discussion.....	22
Conclusion.....	24

Recommendations.....	25
REFERENCES.....	26
APPENDICES.....	29
Appendix A: Student Questionnaire.....	30
Appendix B: Instructor Interview.....	36
Appendix C: Instructor Demographic Charts.....	44
Appendix D: Student Demographic Charts.....	50
VITA.....	56

Chapter I

INTRODUCTION

At the end of most higher education classes, students are asked to complete an evaluation of the instructor who taught the course. At The University of Tennessee (UT), the faculty senate requires all classes with five or more students to evaluate the instructor of a given course. There are eleven different evaluation forms that are used in order to assess instructors on the UT campus. The Student Assessment of Instruction System (SAIS) processes these evaluations forms, forwards them to the respective departments, and then publishes them in Tennessee 101 for public use (Guidelines for In-Class SAIS Administration).

Statement of the Problem

Many questions and issues arise concerning students' evaluation of higher education instructors. It is questionable whether a student can provide an unbiased evaluation of a higher education instructor. Students have attitudes toward completing these evaluations that may be positive or negative. Depending on the attitude or thought process behind the evaluation and the students' ability to provide an unbiased opinion, these evaluations may not provide accurate information about the course and instructor being evaluated. If the opinions provided are biased or influenced by external elements, problems may arise concerning the validity of the evaluations. Instructors should not have inaccurate information used against them to determine important decisions such as raises, promotion, and tenure.

Purpose of the Study

The purpose of this study was to examine the validity of student evaluations of selected higher education instructors at UT during the spring and summer semesters of 2005. This study specifically explored issues such as students' ability to evaluate a higher education teacher without bias, faculty members' attitudes toward their evaluations, how evaluations are used, and the consequences that arise from these issues.

Research Objectives

This study had four main objectives:

1. To determine whether students were likely to provide unbiased evaluations of higher education instructors.
2. To determine the attitudes surrounding student evaluations of higher education instructors.
3. To determine whether student evaluations of higher education instructors were a valid form of assessing an instructor.
4. To determine whether higher education instructors take their evaluations seriously and use them as a tool to improve their instruction.

Research Questions

Based on the objectives listed, the primary research questions were:

1. What types of characteristics do students consider when completing an evaluation of a higher education instructor?
2. How serious are students about evaluations of higher education instructors?
3. How do gender, age, student status, or average grades make differences in how a student evaluates a higher education instructor?

4. How seriously do higher education instructors take their student evaluations?
5. How do higher education instructors use student evaluations of their performance and class to improve their teaching and classes?

Assumptions

When administering a self-report questionnaire about behaviors, attitudes, perceptions, and preferences, one assumption is that the participants will provide accurate information. The students and instructors in this study may respond as they think they should instead of responding with accurate behaviors, attitudes, perceptions and preferences. Since the student questionnaire and the instructor interview is a self-report, it was assumed that students and instructors responded to questions accurately.

Limitations

This study was limited by the instructors' agreement to be interviewed and to allow the use of their class time in which to administer student questionnaires. The instructors were allowed to choose the course to be used in this study for their convenience and available class time, therefore limiting the findings. Some departments were not able to participate because of lack of available time during class.

Definition of Terms

Sexy-horrible teacher- An instructor that a student finds sexy but who was not an effective teacher.

Chapter II

REVIEW OF LITERATURE

According to author Paul Trout, the majority of higher education institutions in the United States require students to complete evaluations of their instructors (Trout, 2000). Trout believes that student evaluations of college instructors are more like a “customer satisfaction survey” (Trout, 2000) and do not measure the actual performance of the instructor. Trout’s concern is that if student evaluations measure satisfaction instead of performance of the instructor and they are then used to determine teachers benefits (such as raises and tenure), it will drive instructors to do whatever it takes to get positive evaluations in order to get raises and other perks. This would mean higher education instructors would alter their classes to satisfy the student or administration. Instructors (intentionally or unintentionally) would grade easier, reduce requirements of courses, and give easier tests to satisfy their students. If this is true, the American higher education classroom is in great danger of providing what the author refers to as “dumbed-down education” (Trout, 2000).

Nasser and Fresko (2002) conducted a study of 170 faculty members using questionnaires to evaluate attitudes toward student evaluations of college instructors. In this study, 55.8% of the respondents believed that students take evaluations seriously, and 43.7% believed that evaluations reflected quality of instruction. As well, 57 % believed that an instructor had to be a good instructor to get a good evaluation. There were mixed feelings about whether evaluations should be used as a factor in granting tenure and promotion (Nasser and Fresko, 2002). Results also showed that 71% of these participants

agreed that students' evaluations help instructors to improve their teaching skills in some way, and 58.1% agreed that they help pinpoint weak areas in the course, yet only 2.5% to 10.3% reported making any changes to their courses as a result of students' evaluations (Nasser and Fresko, 2002). Also evaluated in this study were the factors that may have influenced the feelings of the respondents. It was found that male instructors were more likely than female instructors to make changes in their courses due to students' evaluations.

There have been a number of studies conducted on the attitude of the students concerning their evaluations of higher education instructors and issues surrounding them. Kenneth Feldman (1976) conducted a research study on students' opinion of the most favorable characteristics of the college instructor. He found a large body of literature that established that students consider quite a few characteristics associated with effective teaching. Some of these characteristics include knowledge of the subject matter, students' ability to understand the teacher, enthusiasm, helpfulness, and openness (Feldman, 1976). Fascinatingly enough, Feldman also discovered that even though these students viewed these characteristics as important, they didn't necessarily use them to assess their instructors on evaluations.

So, knowing that in general students seem to have a set of characteristics they feel are important for a good instructor to possess; but at the same time knowing they don't always use these characteristics to evaluate their instructors leaves one to ponder what fills in this gap between the characteristics of a good higher education teacher and a higher education instructor who receives a positive student evaluation at the end of the semester.

There were two studies conducted that help to fill in this gap between the characteristics of a good instructor and the instructor who receives positive evaluations. Felton, Mitchell, and Stinson (2004) conducted an attention-grabbing study to evaluate whether college students are biased when evaluating college professors. They used a website popular with college students to evaluate whether college students showed bias in areas of “perceived quality, easiness, and sexiness” (Felton et al, p. 91). The authors concluded that students who used this medium to evaluate teachers were significantly biased by two factors: the perceived sexiness of the teacher as well as the easiness of the course that was taught.

Another study performed by Best and Addison (2000) explored the relationship between “the perceived warmth” (Best and Addison, p.60) of an instructor and the students’ evaluation of that instructor. Best and Addison concluded that college students did give higher evaluations if they perceived their instructors as being warm. This study was performed over three semesters in which the senior author taught psychology classes in different ways in order to see how her evaluations were effected.

Spencer and Schmelkin (2002) completed a comprehensive opinion survey of students and their various perspectives on evaluations of college instructors and how they should be used. This particular group of students felt comfortable completing student evaluations of their instructors and did not mind completing them. However, they did feel that their opinions did not count for much. These students most often felt that instructor evaluations should be used to provide information to the instructor in order to improve his/her instruction. Most of the students in this study reported not consulting published results of instructor evaluations for any reason.

Filak and Sheldon (2003) used a theory of self-determination that assumes that human beings psychologically need “autonomy, competence, and relatedness” (Filak & Sheldon, p. 236) to conduct two studies that examined how students’ psychological needs predicted university instructor evaluations. The authors predicted that when students’ psychological needs (autonomy, competence, and relatedness) were met, both course and instructor evaluations would be higher (Filak and Sheldon, 2003). The first study examined the force of other factors such as class size and topic being taught, whereas the second study zeroed in on characteristics of the instructors. The authors revealed that the results of both studies showed that all three needs (autonomy, competence, and relatedness) were predictors of instructor evaluations. However, only autonomy and competence were shown to predict positive course evaluations (Feliak and Sheldon, 2003). Other relevant findings included were in the second study, in which the authors found that if instructors taught classes repeatedly (to different groups of students), their ratings tended to be lower. This might suggest that the psychological needs of the students were no longer being fulfilled if the instructors taught sequential classes (Feliak and Sheldon, 2003).

Yet another study conducted by Williams and Ceci (1997) evaluated the outcome of teaching two identical courses. The same instructor taught courses with a minor adjustment to the second course; he taught it with more “enthusiasm” (Williams and Ceci, p. 15). His evaluations in the second course were not only higher for enthusiasm but also higher for other characteristics as well. In other words, because he was more enthusiastic, students perceived him as being more knowledgeable, organized, and available to help them (Williams and Ceci, 1997).

However, when “push comes to shove”, most people would agree that one important thing in the classroom is that students learn something. Arthur, Tubre, Paul, and Edens (2003) conducted a study using a pre-test and post-test design to measure how much the students had learned over the semester. These authors concluded that students who learned the most as determined on the pre- and post-tests over the semester received the highest grades in the course. There was also a small relationship between instructor evaluations and how much the student learned in the class (as measured by post- and pre-tests) (Arthur et al, 2003).

Another problem with the validity of student evaluations of higher education instructors is the instructors’ leniency (coursework and grades). Marsh and Roche (2000) examined how grading leniency and course workload affected instructor evaluations. Surprisingly, they concluded that instructors who give higher grades and less workload do not get higher evaluations from students. These authors concluded that a high instructor evaluation could be achieved through presenting demanding material that would challenge that student (Marsh and Roche, 2000).

Yunker and Yunker (2003) did a study comparing student achievement and their evaluations of higher education instructors in response to the concern that grade inflation may be a result of leniency of higher education instructors expecting a better evaluation from students. Tracking students in business classes, Yunker and Yunker (2003) found that the students who came from classes in which the instructor was rated high did not do as well when they advanced to the next level course.

Shevlin, Banyard, Davies, and Griffiths (2000) used questionnaires distributed to students to determine a correlation between students' perceived charisma of a lecturer and teaching effectiveness. Shevlin et al (2000) found an association between whether the students viewed the instructor as charismatic and how they viewed their teaching ability.

There are many factors that determine how a student evaluates a higher education instructor. It seems as though perceived teaching effectiveness is often evaluated by personality characteristics of the instructor. Kenneth Feldman (1986) conducted a review of literature to examine how perceived effectiveness of college instructors relates to the instructors' personality characteristics. Feldman (1986) found a positive correlation between perceived instructor effectiveness and positive self-regard, energy, enthusiasm, and positive regard for others. The types of traits that students seemed to like in instructors and also tended to give higher ratings to were the exhibited ability to control the class, reflectiveness, flexibility, emotional stability, sociability, responsibility, and brightness. The one trait that the students did not like and tended to rate lower was anxiety or neuroticism (Feldman, 1986).

It seems as though there are many factors that influence the student evaluation of a higher education instructor. Feldman (1976) found that students did have a set of favorable characteristics to define a good instructor. However, these characteristics were not necessarily used to evaluate an instructor. Feldman (1986) found that students had certain personality traits of instructors that they liked and rated those instructors who possessed those traits higher on evaluations. The question remains whether personality traits necessarily define a good or a bad teacher. Shevlin et al (2000) found an association to charisma and perceived teaching ability. When Williams and Ceci (1997) conducted a

study on enthusiasm, they found that students perceived an instructor as better just because he taught with more enthusiasm. Felton et al (2004) found that students rated instructors they perceived as easy (in relation to the course) or sexy (in relation to appearance) higher than those they didn't perceive as easy or sexy.

Summary of Literature

According to this review of literature, students have a set of characteristics that they deem effective higher education instructors to possess. Rather than use these characteristics to complete an evaluation, it seems as though students are more likely to evaluate instructors on personal characteristics such as appearance, enthusiasm, easiness, and charisma, which do not necessarily define an instructor's ability to teach effectively. This being true, the validity of such evaluations is compromised because the purpose of evaluations is to rate instructors' teaching abilities, not their personal characteristics.

Chapter III

METHODS

The methods used in this research report include both student questionnaires and instructor interviews (Appendix A & Appendix B) collected at UT. These questionnaires were obtained in order to examine the validity of students' evaluations of a higher education instructor at the university and to gain insight into the attitudes surrounding these evaluations.

A variety of higher education instructors were interviewed and students in one of their classes completed a questionnaire. For the instructors' convenience, they chose the class to evaluate based on available class time. Average evaluation scores were also used for each instructor.

Participants

Participants in this study were limited to a sample of higher education instructors at UT who agreed to be interviewed and who allowed class time to administer the student questionnaire in one of their classes. In order to obtain a random sample of participants, classes were selected from a variety of academic departments on the UT campus. Child and Family Studies, Instructional Technology and Educational Studies, Engineering, Sports Studies, and the School of Architecture were selected. Eight instructors were chosen and agreed to participate. The instructors ranged in teaching experience and status (see Appendix C for a chart of demographic information on the instructors).

Four graduate classes and four undergraduate classes were selected by the instructor to administer student questionnaires (see Appendix D for a chart of student demographic information).

Implementation

The higher education instructor interviews and the student questionnaires were conducted during the spring and summer sessions of 2005. Interviews were conducted at the instructors' convenience, and the student questionnaires were completed during a class period allocated by the instructor.

Data Analysis

The data were entered and analyzed using SPSS. Simple descriptive analysis methods such as Pearson Correlations were performed and used to evaluate these data. Current instructor evaluations were averaged and compiled into a score based on a 5 point scale (5 being highest possible score). Because UT uses multiple forms in which to evaluate a higher education instructor, only like categories were used to obtain the average score.

CHAPTER IV

RESULTS

Instructor Interviews

The instructors interviewed for this study agreed that they used student evaluations of their performance in some form. Two instructors (25%) said that they used evaluations of their performance to figure out how to get higher scores on their evaluations.

The instructors in this study all agreed that student evaluations of higher education instructors should be used to improve the overall effectiveness of teaching. Two instructors (25%) replied that evaluations should be used to determine salaries, raises, and tenure of instructors. Two instructors (25%) felt that evaluations should be used to get a sense of student expectations for the instructor and the class.

The majority of instructors (87.5%) felt that student evaluations of higher education instructors should be published. Two instructors (25%) felt that students should not have access to evaluation results, and one (12.5%) felt that administrators should not have access to instructor evaluations. Five instructors (62.5%) felt that instructor evaluations should not be published for the public.

Seven instructors (87.5%) replied that they took student evaluations of their performance seriously. Six instructors (75%) felt that students were not able to provide unbiased evaluations of their performance. The instructors had a difference of opinion when asked to respond whether they thought instructors received fair evaluations of their performance from students. Four (50%) thought instructors do receive fair evaluations

and four (50%) believed that instructors did not receive fair evaluations, of their performance.

When the instructors were asked if students use factors unrelated to the overall effectiveness of teaching to evaluate an instructor, all but one (12.5%) responded that they felt students do use factors unrelated to the overall effectiveness of teaching in order to evaluate an instructor. All eight instructors agreed that students use personal characteristics of the instructor in evaluation of instructors. Six instructors (75%) felt that students use personal feelings toward an instructor to base an evaluation. Seven instructors (87.5%) replied that students used perceived grades to evaluate instructors.

The instructors in this study rated personality characteristics in relation to their importance. Most instructors (87.5%) indicated that availability to help, knowledge of subject matter, and preparedness were very important characteristics of a higher education instructor. Two instructors (25%) responded that openness to others' views was not important and that the ability to stimulate interest was only somewhat important. Four instructors (50%) thought that enthusiasm was only somewhat important. The instructors had a wide range of opinions when it came to friendliness. Two instructors (25%) felt that friendliness was very important, and two other instructors (25%) felt that it was not important at all. All other instructors felt that friendliness was somewhat important.

Pearson correlations that were statistically significant were found between the following items: a positive correlation between the fairness rank and the perceived D's & F's reported on teacher evaluations [$r = .855$, $p < .01$, $n = 8$]. The higher the fairness rank, the less likely it was for the teacher to have perceived D's & F's as reported by students on evaluations. There was also a correlation between the status of the instructor and

perceived D's & F's on evaluations [$r = .799, p < .05, n = 8$]. The lower the status of the instructor, the more likely they were to have perceived D's & F's on student evaluations. The number of perceived C's on teacher evaluations was also correlated with the class being taught [$r = .792, p < .05, n = 8$]. Architecture and Engineering classes were more likely to report perceived C's than ITES, CFS, or Sports Studies classes. If instructors reported taking evaluations seriously, they were more likely to report they felt that evaluations should be published and that administrators should have access to the results of student evaluations [$r = 1.000, p < .01, n = 8$]. When instructors reported they felt students should have access to the results of student evaluations, the more likely they were to rate openness to others' views as very important [$r = -.926, p < .01, n = 8$]. There was a correlation as well between the teaching experience and how high the instructor ranked openness to others' views [$r = -.870, p < .01, n = 8$]. Similarly, instructors with more teaching experience tended to rate openness to others' views as very important [$r = -.827, p < .05, n = 8$]. The female instructors were more likely to rate friendliness as being very important [$r = -.745, p < .05, n = 8$].

Student Questionnaires

Students were asked to respond to the factors they use to determine evaluation scores. When students were asked about personal characteristics, 126 (67.7%) replied that they used personal characteristics to rate a teacher on evaluations. When it comes to course rigor, 120 (64.5%) reported that they did use the rigor of the course in order to determine scores on evaluations of their instructors. Seventy-eight percent of students reported that they used how much they learned in the course to rate an instructor on an evaluation. The personal feelings toward the instructor were reported by 119 students

(64%) to be used in order to evaluate an instructor. The perceived grade in the class was reported to be used to evaluate an instructor by 95 students (51.1%). Physical appearance of the instructor was reported to be used to evaluate by 21 students (11.3%). Some students wrote in additional factors that they used in order to evaluate an instructor. There were 18 students (9.7%) who added teaching methods as a factor when evaluating an instructor. Sixteen students (8.6%) wrote in that they used whether or not the instructor was easy to understand as a determinate of scores on instructor evaluations. The instructors' willingness to help students was also written in by 16 (8.6%) of the students as being a factor in determining evaluation scores.

When asked to respond to how student evaluations should be used, 162 students (87.1%) replied that evaluation should be used to improve the overall effectiveness of the instructor. Seventy-nine percent (147) of students felt evaluations should be used to improve the overall effectiveness of the course. Fifty-three students (28.5%) reported feeling that instructor evaluations should be used to determine salaries, raises, and tenure. There were 30 students (16.1%) who wrote in that evaluations should be used to hold instructors accountable for their teaching.

There were 149 students (80.1%) who reported that student evaluations of instructors should be published. When asked who should have access to the results of student evaluations of higher education instructors, 156 students (83.9%) agreed that instructors should have access to the results of student evaluations of instructors. There were 159 students (85.5%) who reported they felt that students should have access to instructor evaluations, as well as 161 students (86.6%) who reported that administrators

should have access to instructor evaluations. Furthermore, 66 students (35.5%) responded that the public should have access to student evaluations of instructors.

When students were questioned about the personality characteristics of a university instructor, 159 students (85.5%) answered that fairness was a very important characteristic of an instructor. As well, 134 (72%) reported that preparedness was a very important characteristic of an instructor. Eighty-six students (46.2%) agreed that friendliness is a very important characteristic of a higher education instructor.

Students were asked to determine how they would rate an excellent teacher that they didn't like. There were 84 students (45.2%) who responded that they would give positive responses to an excellent teacher that they didn't like. Fourteen students (7.5%) said they would give an excellent teacher they didn't like negative responses on evaluations.

Students were asked to determine how they would rate an "everyone gets an A teacher" even though they didn't learn a thing. There were eighty-two students (44.1%) who responded that they would give this type of teacher neutral responses. Fifty-two students (28%) replied that they would give this teacher positive responses, and 43 students (23.1%) said they would give this teacher negative responses. Seven students (3.8%) were not sure how they would rate an "everyone gets an A" teacher.

Twenty-seven students (14.5%) reported that they would give a sexy-horrible teacher positive responses. There were 104 students (55.9%) who replied they would give this type of teacher negative responses.

The students were asked about a horrible teacher that they liked. Neutral responses were most often picked by 88 students (47.3%). There were 34 students

(18.3%) who replied that they would give a horrible teacher that they liked positive responses on an evaluation.

The students were asked to write in some of the traits or characteristics of their best higher education instructor. The most frequent response was knowledge or intelligence ($n = 82$, 44.1%). Fifty students (26.9%) replied that helpfulness was a trait of one of their best higher education instructors.

When students were asked to list traits or characteristics of their worst higher education instructor, 43 students (23.1%) said that a bad personality was a characteristic of their worst higher education instructor. There were 34 students (18.3%) who replied that a characteristic of their worst instructor was the inability to speak English.

There were 146 students (78.5%) who reported taking their evaluations of instructors seriously. There were 45 students (25.8%) who indicated that they knew how evaluations were used at UT.

Several correlations that were statistically significant were found between variables on the student questionnaires. The students who reported the highest average grades were linked to responding that they used how much they learned in a class to base their opinions on evaluations [$r = -.183$, $p < .05$, $n = 186$]. Also correlated with the students who reported the highest average grades were negative reports for the sexy teacher [$r = .350$, $p < .01$, $n = 183$] as well as negative reports for the horrible teacher [$r = -.292$, $p < .01$, $n = 184$]. The students who reported the highest grades were also more likely to take evaluations of higher education instructors seriously [$r = -.300$, $p < .01$, $n = 186$].

The students who reported lower grades were more likely to use the physical appearance of the teacher on which to base their decisions on evaluations [$r = .258, p < .01, n = 186$] as well as taking their personal feelings for an instructor into consideration when evaluating an instructor [$r = .190, p < .01, n = 186$] and using their perceived grade in a course to base evaluations of the instructor [$r = .147, p < .05, n = 186$]. Students who received lower grades also thought that evaluations should be used to determine salaries, raises, and tenure [$r = .200, p < .01, n = 186$] as well as accountability of the instructor [$r = .282, p < .01, n = 186$].

The older the student reported being, the less likely they were to base evaluations on personal feelings of the instructor [$r = -.159, p < .05, n = 186$]. The older the student, the more likely they reported openness to others' views as a very important characteristic of a higher education instructor [$r = -.151, p < .05, n = 185$].

The students' gender was also associated with other factors on this questionnaire. Males were the most likely to write in that teaching methods are used to base opinions when completing evaluations of instructors [$r = .175, p < .05, n = 186$]. More males than females reported that they thought evaluations should be used to improve the overall effectiveness of instructors [$r = .169, p < .05, n = 186$]. More females than males believed that openness to others' views was a very important characteristic of higher education instructors [$r = -.202, p < .01, n = 185$]. More males than females reported that they would give a positive response to sexy instructors on their evaluations [$r = .286, p < .01, n = 183$]. More males also reported that caring about students was a characteristic of one of their best university instructors [$r = .149, p < .05, n = 186$]. More females than males rated knowledge or intelligence in an instructor as higher [$r = -.171, p < .05, n = 186$].

Males also added that being boring [$r = .175$, $p < .05$, $n = 186$] and bad grading [$r = .186$, $p < .05$, $n = 186$] were two characteristics of their worst university instructor.

CHAPTER V

SUMMARY, DISCUSSION, CONCLUSION, AND

RECOMMENDATIONS

Summary

The majority of both students and instructors agreed that student evaluations should be published. As to who should have access to student evaluations of higher education instructors, the majority of the instructors and the students replied that instructors, students, and administrators should have access to evaluations but that the public should not have access to them.

The majority of students agreed that they take evaluations of higher education instructors seriously. A large number of the instructors replied that they took evaluations of their performance seriously. However, very few instructors indicated that students were able to provide unbiased evaluations of their performance. Instructors were at odds about receiving fair evaluations from students. Almost all of instructors thought the students use factors unrelated to the overall effectiveness of teaching to evaluate instructors.

Students and instructors basically agreed that evaluations should be used to improve the overall effectiveness of the instructor and the course; they also believed that evaluations should not be used to determine salaries, raises, and tenure of the instructors.

On the one hand, all instructors agreed that students use personal characteristics of the teacher as a factor to complete evaluations. On the other hand, just a little over half of the students reported using personal characteristics to evaluate instructors. Very few

students reported using the physical appearance of an instructor to base their evaluation of the instructor, but half of the instructors felt that a great number of students use physical appearance to base their opinions of instructors. Just a little over half of the students thought they used the perceived grade in the class as a factor to rate the instructor, but the majority of instructors said they felt that students do use their perceived grade to evaluate an instructor. The fact that a student receives a good or poor grade in class does not necessarily measure the effectiveness of the instructor.

Discussion

Based on the data analysis of this study, the majority of students reported using the personal characteristics of an instructor, the rigor of the course, how much they learned in the course, and their personal feelings toward the instructor to evaluate higher education instructors. Most students agreed that fairness, knowledge of subject matter, and preparedness are very important characteristics of instructors. The goal of student evaluations of higher education teachers is to measure how well an instructor performs in the classroom.

A great number of students reported taking evaluations of higher education instructors seriously. However, very few students replied that they knew how instructor evaluations were used at UT.

Female students in this study felt that openness to others' view was a very important characteristic of higher education instructors; females rated knowledge or intelligence of the instructor higher than males. Male students who participated thought that student evaluations should be used to improve the overall effectiveness of the instructor, and they reported using teaching methods in which to evaluate an instructor.

More male students also reported that they would give a sexy teacher positive responses on an evaluation regardless of the fact they didn't learn a thing in the class. More males replied that being boring and grading poorly were characteristics of the worst higher education instructors and listing how much the instructor cared for students as a characteristic of the best instructors.

The older the student in this study, the less likely they were to report basing an evaluation of a higher education teacher on personal feelings of the instructor. Older students most often rated openness to others' views higher than younger students.

Students in this study who reported receiving higher grades more often reported being serious about the evaluations they complete for their instructors. These students also tended to report not using physical appearance of the instructor to base their opinions on evaluations. While higher grade students used factors such as how much they learned in a class, lower grade students used factors such as their personal feelings for an instructor, physical appearance of the instructor, and their grade in order to evaluate an instructor. The lower grade students more often felt that evaluations should be used in order to determine salaries, raises, and tenure of the instructor.

The majority of instructors felt that students are not able to provide unbiased evaluations and that they use factors unrelated to the overall effectiveness of teaching to base their evaluations. Very few instructors said they do not take student evaluations of their performance seriously.

All the instructors interviewed reported that they use the results of student evaluations in some manner. The majority reported using evaluations to improve the

overall effectiveness as a teacher. Very few instructors said that they use student evaluations to figure out how to receive higher evaluation scores.

Conclusion

Based on the results of this study, students are unlikely to provide unbiased evaluations of higher education instructors. More than half of the students in this study said they use their personal feelings to evaluate their instructors. This indicates that instructors are evaluated based on personal preferences rather than instructional effectiveness. In addition, students reported using personal characteristics of an instructor to determine responses on evaluations, which also measures something other than instructional effectiveness. Students also reported using the perceived grade in the class to base their evaluations of their instructors. It is hard to measure an effective instructor when there is not a consensus for guidelines of an effective instructor. This leaves individual opinions for what makes an effective instructor.

Attitudes measured in this study seemed to be positive for the most part. The majority of instructors as well as the students reported taking instructor evaluations seriously. However, a few instructors reported using the results of evaluations to figure out how to receive higher scores on evaluations. This particular type of instructor would be catering to students' every wish without regard for the task at hand: effective teaching.

Although student evaluations of higher education instructors can be used as an evaluation tool, it is uncertain if it is feasible to consider evaluations a valid form of assessing an instructor. Because there are varying opinions on what constitutes a good or even an effective instructor, it makes it hard to determine just what student evaluations measure. If factors such as the personal feelings toward the instructor are used to

determine the score on an evaluation, this can't possibly measure the instructors' ability to teach. Student evaluations of higher education instructors can be and should be a useful tool to maintain and to improve on the quality of higher education.

Recommendations

Research for further study would be to obtain a bigger sample of higher education instructors. It would be interesting to include not only instructor evaluations but also the student evaluation system (grades) in order to examine the correlation between the two variables. Future research may focus on how students' final grades sway the outcome of higher education instructors' evaluations. Furthermore, it would be noteworthy to look at how the university uses the higher education instructors' evaluation in determining promotion and salary schedule. This research study was only able to accommodate a small portion of the instructors' evaluation process. Finally, a more in-depth study of this subject matter is desirable to find out precisely how important evaluations are to the university as an educational institution and what effects they have on our education system as a whole.

REFERENCES

References

- Arthur, W., Tubre, T., Paulm D. S., & Edens, P. S. (2003). Teaching effectiveness: The relationship between reaction and learning evaluation criteria. *Educational Psychology*, 23 (3), 275-285. **Retrieved October 30, 2004 from Academic Search Premiere Database.**
- Best, J. B. & Addison, W. E. (2000). A preliminary study of perceived warmth of professor and student evaluations. *Teaching of Psychology*, 27 (1), 60-62.
- Feldman, K.A. (1976). The superior college teacher from the students' view. *Research in Higher Education*, 5, 243-288.
- Feldman, K. A. (1986). The perceived instructional effectiveness of college teachers as related to their personality and attitudinal characteristics: A review and synthesis. *Research in Higher Education*, 24 (2), 139-173.
- Felton, J., Mitchell, J., & Stinson, M. (2004). Web-based student evaluations of professors: The relations between perceived quality, easiness, and sexiness. *Assessment & Evaluation in Higher Education*, 29 (1), 91-108. **Retrieved October 30, 2004 from Academic Search Premiere Database.**
- Filak, V. F. & Sheldon, K. M. (2003). Student psychological need satisfaction and college teacher-course evaluations. *Educational Psychology*, 23 (3), 235-247. **Retrieved October 30, 2004 from Academic Search Premiere Database.**
- Guidelines for In-Class SAIS Administration. **Retrieved March 5, 2005 from** <http://oira.tennessee.edu/sais/guide.html>

- Marsh, H. W. & Roche, L. A. (2000). Effects of grading leniency and low workload on students' evaluations of teaching: Popular myth, bias, validity, or innocent bystanders? *Journal of Educational Psychology*, 92 (1), 202-228.
- Nasser, F. & Fresko, B. (2002). Faculty views of student evaluation of college teaching. *Assessment & Evaluation in Higher Education*, 27 (2), 187-198. **Retrieved October 30, 2004 from Academic Search Premiere Database.**
- Shevlin, M., Banyard, P., Davies, M., & Griffiths, M. (2000). The validity of student evaluation of teaching in higher education: Love me, love my lectures? *Assessment & Evaluation in Higher Education*, 25 (4), 397-405. **Retrieved October 30, 2004 from Academic Search Premiere Database.**
- Spencer, K. J. & Schmelkin, L. P. (2002). Student perspectives on teaching and its evaluation. *Assessment & Evaluation in Higher Education*, 27 (5), 397-409. **Retrieved October 30, 2004 from Academic Search Premiere Database.**
- Trout, P. (2000). Flunking the test: The dismal record of student evaluations. *Academe*, 86 (4), 58-61.
- Williams, W. M. & Ceci, S. J. (1997). "How'm I doing?": Problems with student ratings of instructors and courses. *Change*, 29 (5), 13-23.
- 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.

APPENDICES

APPENDIX A

Student Questionnaire

February 13, 2005

Dear Research Participant,

My name is Kristy McRiffey and I am a graduate student in the College of Education, Health, and Human Sciences. For the purpose of completing my Masters Degree in Curriculum, Research, and Evaluation I am conducting a study examining student evaluations of university instructors. The attached questionnaire is designed to gather your opinions on student evaluations of university teachers.

I would appreciate your completion of this questionnaire. The results of this questionnaire are confidential and you do not need to put your name on it. Your participation in this study is voluntary and you may decline to participate at any time without penalty. The completion and return of this questionnaire implies consent to participate.

Thank you in advance for your participation. If you have any questions about this study, you can contact me via e-mail at kmcriffe@utk.edu.

Sincerely,

Kristy McRiffey

How Do You Rate Those Instructors?

The Validity of Student Evaluations of Higher Education Instructors

Student Questionnaire

Please reply to these questions concerning student evaluations of college instructors. Check the answer that best fits your feelings.

1. When you complete an evaluation of a university teacher (lecturer, professor, or instructor) on what do you base your opinions? (Check all that apply)

- ☐ (A.) Personal characteristics of the teacher
- ☐ (B.) Rigor of the course
- ☐ (C.) How much you learned in the course
- ☐ (D.) Physical appearance of the instructor
- ☐ (E.) Your personal feelings toward the instructor (If you liked them or not)
- ☐ (F.) The grade you perceived you would receive
- ☐ (G.) Some other factor, please specify: _____

2. In your opinion how should student evaluations of university teachers be used? (Check all that apply)

- ☐ (A.) To improve the overall effectiveness of the instructor
- ☐ (B.) To improve the overall effectiveness of the course
- ☐ (C.) To determine salaries, raises, and tenure of the instructor
- ☐ (D.) Other ways, please specify: _____

3. Do you think instructor evaluations should be published? (Please check the best answer)

- ☐ (A.) Yes
- ☐ (B.) No

4. In your opinion who should have access to student evaluations of university instructors? (Please check all that apply)

_____ (1.) Teachers

_____ (2.) Students

_____ (3.) Administrators

_____ (4.) Public

5. Please rate the following traits or personality characteristics of university instructors by checking the **Very Important**, **Somewhat Important** or **Not Important** boxes below, and then rank them in order of importance by assigning them a number and placing it in the **Rank** box. (1= higher, 8= lower)

Trait or Characteristic	Very Important	Somewhat Important	Not Important	Rank
Fairness				
Availability to help				
Knowledge of subject matter				
Openness to others views				
Enthusiasm				
Ability to stimulate interest				
Preparedness				
Friendliness				

6. Please determine how you would rate the following instructors and place a check in the box that best describes how you would rate this instructor.

Type of Teacher	Positive Responses	Neutral Responses	Negative Responses	Not Sure
Excellent teacher that you didn't like				
An everyone gets an "A" teacher –you didn't learn a thing				
A "sexy" one-even though a horrible teacher				
Horrible teacher you liked				

7. What are some traits or characteristics of your best higher education instructors?

8. What are some traits or characteristics of your worst higher education instructors?

9. Do you take evaluations of instructors seriously? (Please check the best answer)

_____ (A.) Yes

_____ (B.) No

10. Do you know how instructor evaluations are used at the university in which you are enrolled? (Please check the best answer)

_____ (A.) Yes

_____ (B.) No

11. I am a(n)

_____ (1.) Undergraduate Student

_____ (2.) Graduate Student

_____ (3.) Other, please explain: _____

12. I am a

_____ (1.) Male

_____ (2.) Female

13. I normally average what type of grades. (Please check the best answer)

_____ (A.) A's and B's

_____ (B.) C's

_____ (C.) D's and F's

14. Your age:

_____ (A.) 18-21 years

_____ (B.) 22-25 years

_____ (C.) 26-30 years

_____ (D.) 30 or above years

APPENDIX B

Instructor Interview

INFORMED CONSENT STATEMENT

How Do You Rate Those Instructors? - The Validity of Student Evaluations of Higher Education Instructors

You have been invited to participate in a research study. The purpose of this study is to examine the validity of student evaluations of higher education instructors. This study specifically explores issues such as students' ability to evaluate a higher education instructor without bias, faculty members' attitudes toward their evaluations, how evaluations are used, and the consequences that arise from these issues. This study has four main objectives: 1) To determine whether students are likely to provide unbiased evaluations of higher education instructors. 2) To explore the attitudes surrounding student evaluations of higher education instructors. 3) To determine whether student evaluations of higher education instructors are a valid form of assessing an instructor. 4) To determine whether higher education instructors take their evaluations seriously and use them as a tool to improve their instruction.

PARTICIPANTS

Participants in this study are limited to a sample of higher education instructors at UTK who agree to be interviewed (requiring about 45 minutes) and who will allow class time to administer the student questionnaire in one of their classes (requires about 15 minutes). In order to obtain a random sample of participants, classes were selected from a variety of academic departments on the UTK campus. The Departments of Child and Family Studies, Instructional Technology and Education Studies, Engineering, Psychology, and the School of Architecture were selected.

RISKS

There are no foreseeable risks involved in any procedures that will be used in this study.

BENEFITS

The benefits achieved in this study will be an understanding of the validity of student evaluations of higher education instructors.

CONFIDENTIALITY

The information obtained in this study will be kept confidential. Data will be stored securely and will be made available only to persons conducting the study unless participants specifically give permission in writing to do otherwise. No reference will be made in oral or written reports which could link participants to the study.

_____ Participant's initials

EMERGENCY MEDICAL TREATMENT

The University of Tennessee does not "automatically" reimburse subjects for medical claims or other compensation. If physical injury is suffered in the course of research, or for more information, please notify the investigator in charge.

CONTACT INFORMATION

If you have questions at any time about the study or the procedures (or you experience adverse effects as a result of participating in this study) you may contact the researcher, Kristy McRiffe, at kmcriffe@utk.edu, or by telephone at (865) 938-8149. If you have questions about your rights as a participant, contact the Office of Research Compliance Officer at (865) 974-3466.

PARTICIPATION

Your participation in this study is voluntary; you may decline to participate without penalty. If you decide to participate, you may withdraw from the study at any time without penalty and without loss of benefits to which you are otherwise entitled. If you withdraw from the study before data collection is completed, your data will be returned to you or destroyed.

CONSENT

I have read the above information. I have received a copy of this form. I agree to participate in this study.

Participant's signature _____ Date _____

Investigator's signature _____ Date _____

How Do You Rate Those Instructors?

The Validity of Student Evaluations of Higher Education Instructors

Instructor Interview Questions

1. How do you use student evaluations of your performance? (Indicate all that apply)

_____ (A.) To improve your overall effectiveness as a teacher.

_____ (B.) To figure out how to best receive high scores on evaluations.

_____ (C.) You do not use student evaluations of your performance.

_____ (D.) Some other way, please specify: _____

2. In your opinion, how should student evaluations of higher education instructors be used? (Indicate all that apply)

_____ (A.) For the instructor's personal use to improve teaching effectiveness.

_____ (B.) For the instructor's personal use to improve class effectiveness.

_____ (C.) To determine salary, raises, and tenure.

_____ (D.) Other way(s), please explain: _____

3. Do you think student evaluations of higher education instructors should be published?

_____ (A.) Yes

_____ (B.) No

4. Do you take students' evaluations of your performance seriously?

_____ (A.) Yes

_____ (B.) No

5. Who should have access to students' evaluations of higher education instructors?
(Indicate all that apply)

_____ (1.) Teachers

_____ (2.) Students

_____ (3.) Administrators

_____ (4.) Public

6. Do you think students are able to provide unbiased evaluations of your performance?

_____ (A.) Yes

_____ (B.) No

7. In general do you think instructors receive fair evaluations of their performance from students?

_____ (A.) Yes

_____ (B.) No

8. Do you think students use factors that are unrelated to overall effectiveness of teaching in which to evaluate instructors?

_____ (A.) Yes

_____ (B.) No

9. In your opinion, what factors, other than overall effectiveness of teaching, do students use to base their evaluations? (Indicate all that apply)

_____ (A.) Personal characteristics of the teacher.

_____ (B.) Rigor of the course.

_____ (C.) Physical appearance of the teacher.

_____ (D.) Their personal feelings toward the teacher.

_____ (E.) The grade they felt they would receive.

_____ (F.) Others, please specify _____

10. Please rate the following traits or personality characteristics of higher education instructors by indicating **Very Important**, **Somewhat Important**, or **Not Important** boxes below, then rank them in order of importance by assigning them a number and placing it in the **Rank** box. (1= higher, 8 = lower)

Trait or characteristic	Very Important	Somewhat Important	Not Important	Rank
Fairness				
Availability to help				
Knowledge of subject matter				
Openness to others' views				
Enthusiasm				
Ability to stimulate interest				
Preparedness				
Friendliness				

11. I am a(n):

_____ (1.) Professor

_____ (2.) Associate Professor

_____ (3.) Assistant Professor

_____ (4.) Lecturer

_____ (5.) Instructor

_____ (6.) Other, please specify: _____

12. I am a

_____ (1.) Male

_____ (2.) Female

13. My teaching experience:

_____ (A.) 5 or fewer years

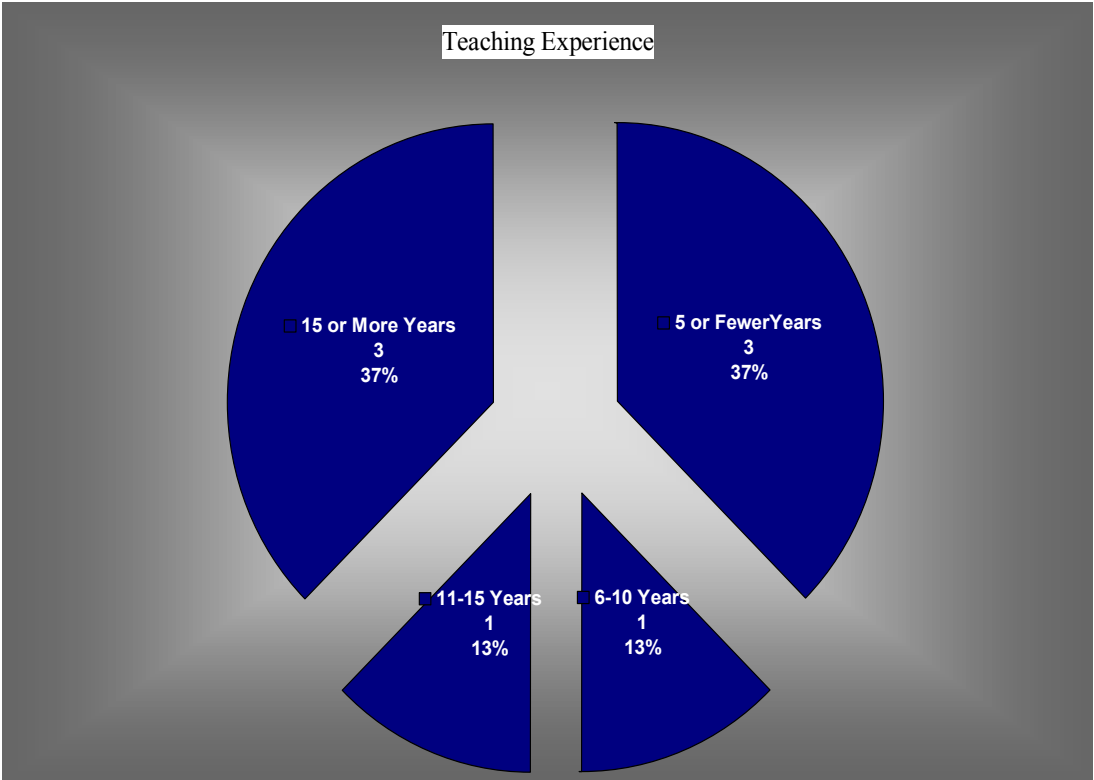
_____ (B.) 6 to 10 years

_____ (C.) 11 to 15 years

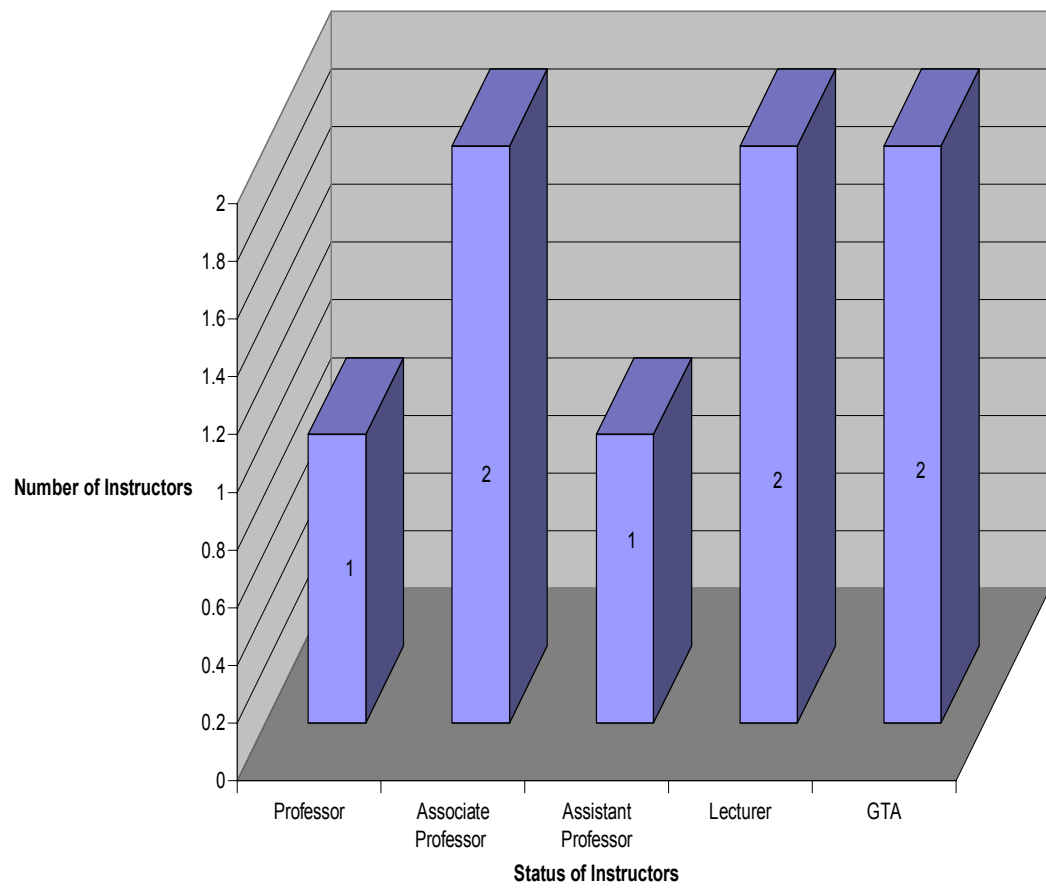
_____ (D.) More than 15 years

APPENDIX C

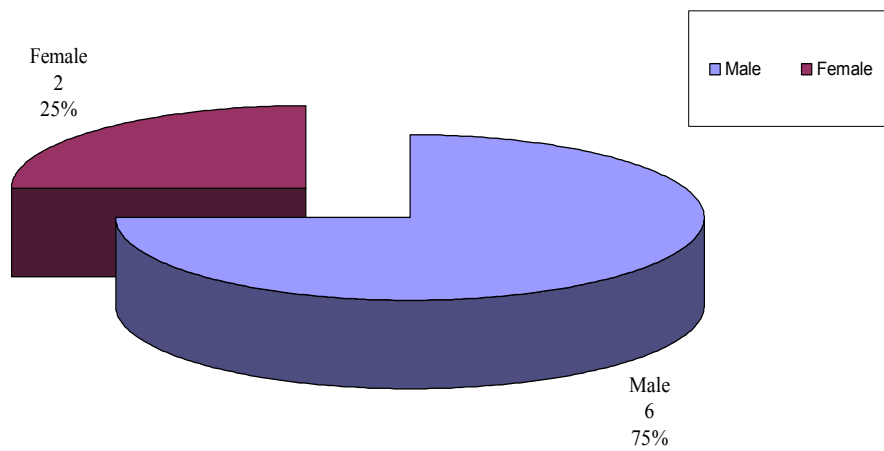
Instructor Demographic Charts



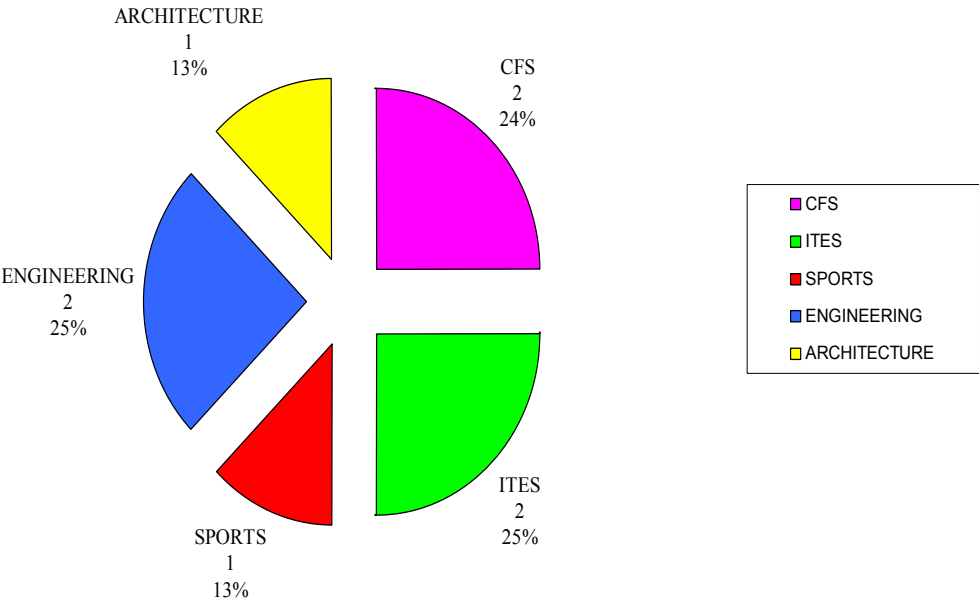
Instructor Status



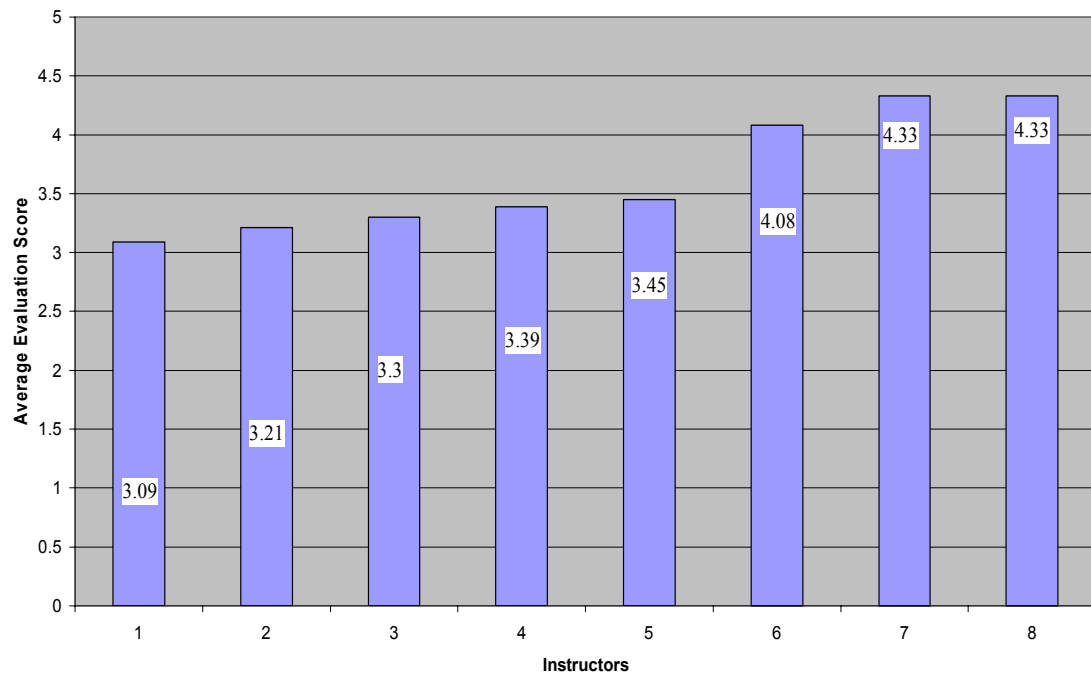
Instructor Gender



Department/College of Instructor



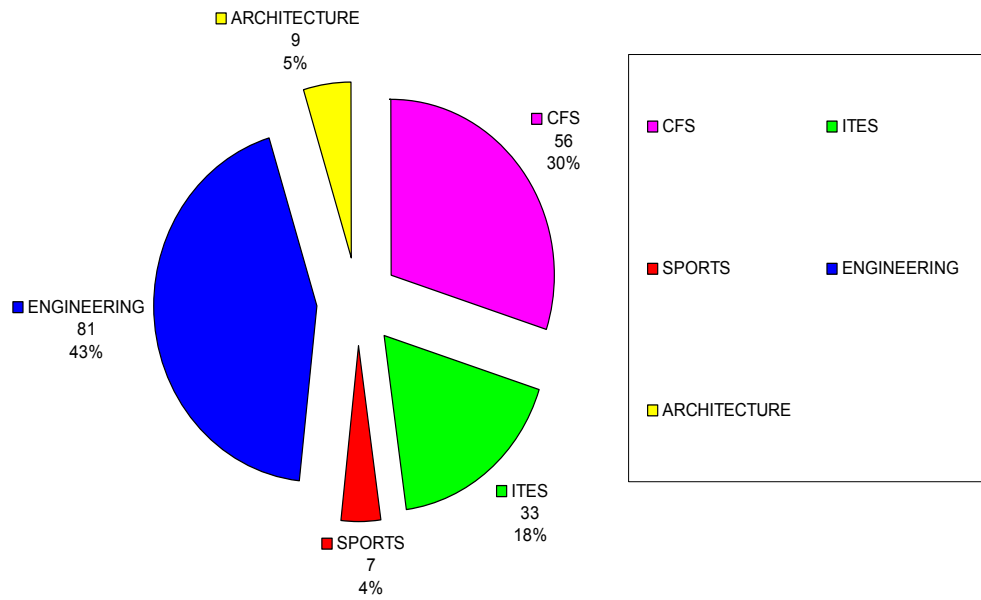
Average Evaluation Scores



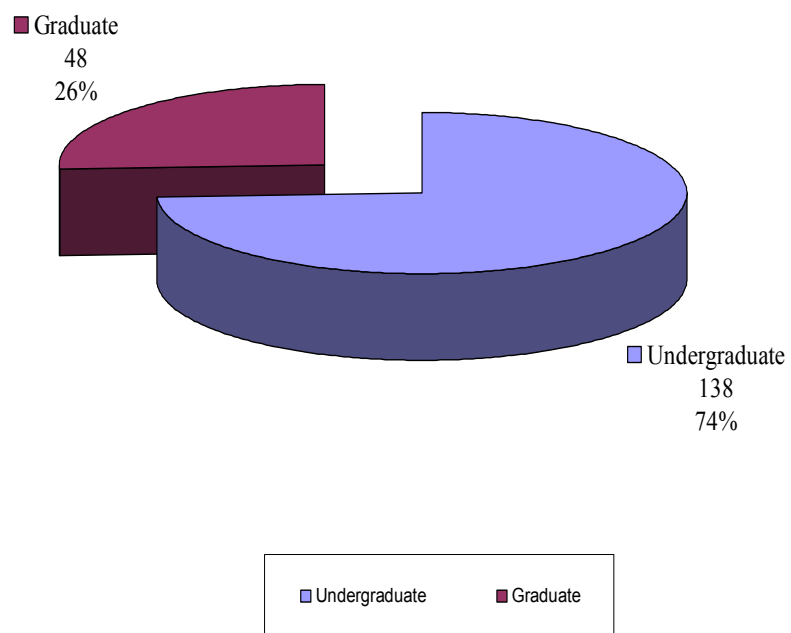
APPENDIX D

Student Demographic Charts

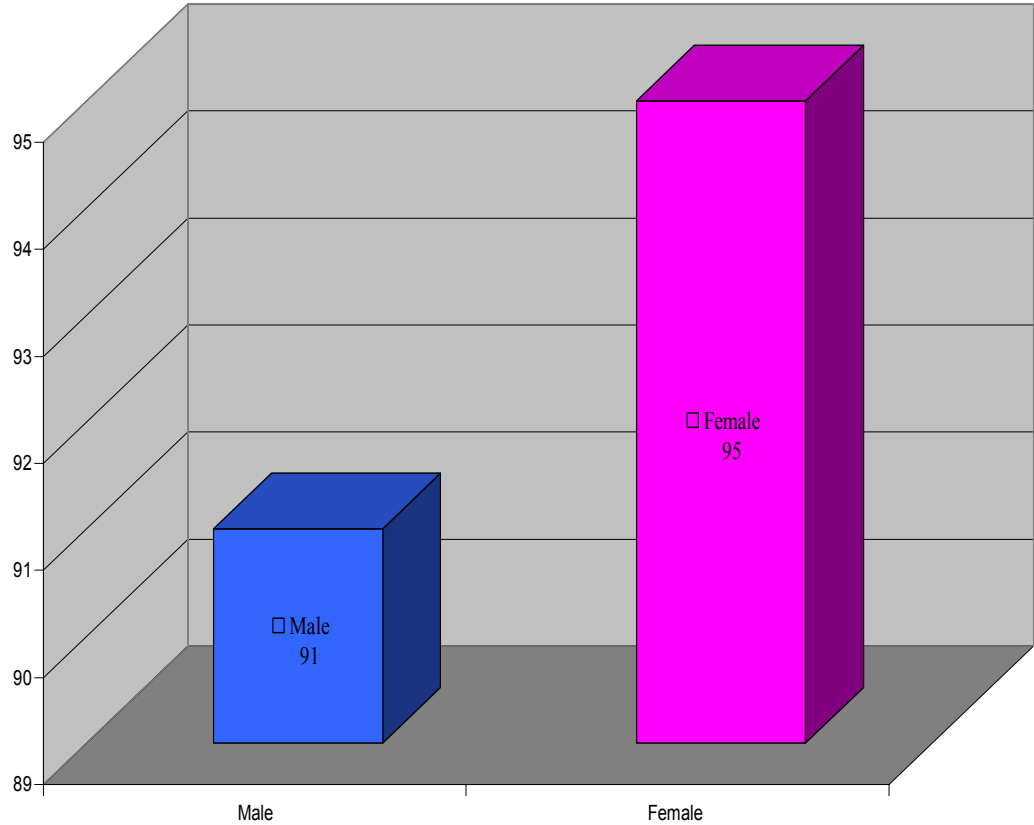
STUDENT MAJOR



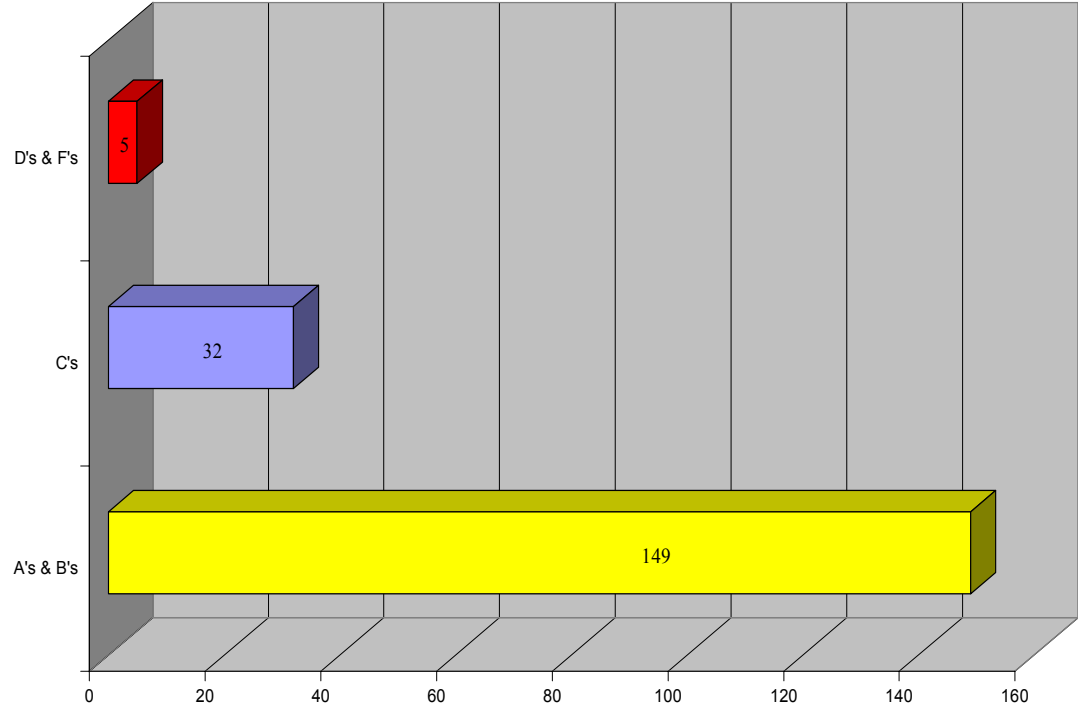
STUDENT STATUS



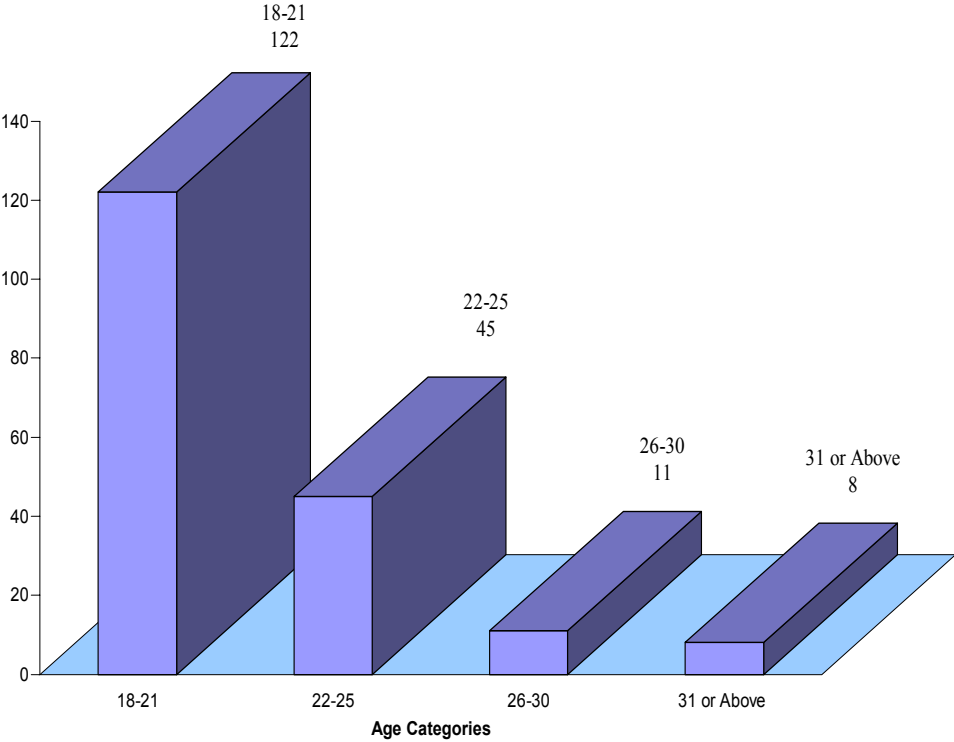
STUDENT GENDER



STUDENT REPORTED GRADES



STUDENT AGE



VITA

Kristy McRiffey was born in Muncie, Indiana, on February 5, 1966. She grew up in Oakville, Indiana, and graduated from Cowan High School in 1984. She attended Ivy Tech, Ball State University, Indiana Wesleyan University, and The University of Tennessee on the road to achieve a Bachelor of Science degree in Human Ecology from UT in May 2003. Kristy became very interested in studying education after completing a research project for the undergraduate research fair in 2003. Kristy has worked for Dr. Cheryl Buehler on The Family Life Project since 2002. She intends to pursue a Ph.D. in Assessment and Evaluation.