## **Student Perceptions of Teaching Evaluations: Continuation**

Linda Ann Carson Lander University 320 Stanley Avenue Greenwood, SC 29649 Meredith Uttley Lander University 320 Stanley Avenue Greenwood, SC 29649

## ABSTRACT

Accountability continues to be an issue in all areas of academia from the accreditation of the institution to student evaluations of their professors. Accountability ratings can impact continued employment, tenure and promotional decisions, and probation for an institution. The goal of this study is to explore student perceptions and attitudes regarding the evaluation of the teaching process and how that may be related to a professor's personality. About three weeks after evaluations were completed, a short survey using Likert-scale questions was distributed to assess students' seriousness and perceptions regarding those student evaluations of teaching. The questionnaire was administered to students in the College of Business and Public Affairs, home to Business, Political Science and Sociology majors. Not surprisingly, students did not take evaluations completed on day two as seriously as those completed on day one of the evaluation period. Fewer students this year than last believe that the instructors took the results of the IDEA evaluation seriously. Over two-thirds of students believe that the IDEA instrument is an accurate means of providing feedback about class teaching objectives to the administration as well as to the professor. In contrast, only slightly more than half of the students perceive that the administration is serious in its use of the evaluation. As expected, students agreed that they are more likely to rate a professor positively if they like the professor, and negatively if they do not like the professor. Student evaluations were also influenced by the subject matter of a class. Slightly less than half of the students believe that the most effective instructors receive the best evaluations, a discouraging result.

### **INTRODUCTION**

Accountability is increasing in all levels of academia from the accreditation of the institution to student evaluations of their professors. Faculty evaluations of their chairs, deans, and the president of the institution are common. Students are now referred to as consumers (Gursoy and Umbreit 2005 and Schnieder et al. 1994) and are being given the power that other types of consumers have when dealing with all types of suppliers (Read, Rama and Raghunandan 2001; Haskell 1997). Student evaluations of teaching (SETs) were originally designed to help instructors improve the quality of their instruction and courses, a formative function (Germain 2005; Birnbaum 1999; Haskell 1997; Rifkin 1995). SETs continue to provide formative feedback for instructors (Aultman 2006), but increasingly they are used by administrators for faculty reappointment, tenure, promotion, and pay increase recommendations, a summative function (Onwuegbuzie et al. 2007; Read, Rama and Raghunandan, 2001; Jackson et al. 1999; Seldin 1999; Pike 1998; Cashin and Downey 1992). Student evaluations are now being used in other countries to assist administrators in their quest for accountability (Burden 2009 and Leite et al. 2006). Student evaluation forms may be designed by individual instructors, by departments, by colleges, or increasingly by outside agencies. Regardless of the instrument used, research in the past, has supported the assumption that students make valid and reliable estimates of their learning (Hoyt and Perera 2000; Cashin 1995; Seldin 1993). SETs should be highly influenced by variables demonstrated to be strongly associated with effective teaching. Unfortunately, researchers have reached no consensus on a definition of quality or effective teaching (Onwuegbuzie et.al. 2007; Germain and Scandura 2005; Okpala and Ellis 2005; Marsh 2001; Jackson, et al 1999; Clayson and Haley 1990). Researchers have addressed student perceptions related to characteristics of effective teaching (Onwuegubzie et al. 2007; Surrant and Desselle 2007; Marsh 2001), dimensions of student perceptions of teaching effectiveness (Brown, 2008; El Hassan 2009; Leite et al. 2006 and Jackson et al. 1999), and student perceptions of learning (Sprinkle 2008;

Addison, Best and Warrington, 2006; Gursoy and Umbreit 2005). Grayson (2004) found that perceived professor performance influenced general student satisfaction with a university program.

Gursoy and Umbreit (2005) and Marsh (2001) explored definitions of good and bad workloads assigned to students. Students define workloads by the amount of time they spend on productive, valuable activities related to the course. The myth is that less workload for students leads to higher student ratings. Gursoy and Umbreit (2005) and Marsh (2001) found that students valued, thus giving positive ratings, for a high good workload and negative ratings for a high bad workload (activities that students did not believe were productive for learning in the course). Of course, there is a point where too much good workload was just too much, thus lowering the rating. Cashin (1995) suggests that evaluations based on workloads support the validity of student ratings, but Cashin does not distinguish between "good or bad" workloads. Lenient grading and its impact on positive ratings has been the focus for a number of researchers, but results have been mixed (Surratt and Desselle 2007; Marsh 2001; Jackson et al. 1999; Cashin 1995). There is a general positive correlation between grades, either expected or actual, and student evaluations of teaching (Nerger et al. 2007; Millea and Grimes 2002; Johnson 2002; Marsh 2001; Greenwald and Gillmore 1997; Jackson et al. 1990; Cohen 1981), but Marsh and Roche (2000) caution that good grades can come from higher motivation and greater interest in the subject matter and need not constitute bias. In fact, "workload, expected grades, and their relations with SETs were stable over 12 years" (Marsh and Roche 2000). However, Addison, Best and Warrington (2006) found that students rated instructors more favorably when a course was easier than expected, but gave a less than favorable evaluation if the course was harder than the student expected.

Age (Sprinkle 2008; Horner, Murray and Rushton 1989), race (Burden 2009; Schulze and Tomal 2006; Anderson and Smith 2005) and gender (Basow 1995; Kierstead,, D'Agostino and Dill 1988) of both the instructor and the student have an impact on the students' perceptions. According to Sprinkler (2008), "As respondents' age increased, they were more likely to believe instructors over age fifty-five were more effective than younger instructors, whereas younger respondents were more likely to state that instructors under fifty-five were more effective." Further, "Older respondents (*non-traditional students*) were more likely to take responsibility for their own learning and grade, rather than place the burden upon the professor/instructor" (Sprinkle 2008).

Schulze and Tomal (2006) determined that there were also perceptual differences among students of professors' competence based on gender and race. According to Kierstead, D'Agostino and Dill (1998), what students expect from professors depends, in part, on students' gender expectations. Females receive positive ratings for smiling while male professors receive negative ratings for smiling. Others have found that the age, race and gender of both instructor and student have shown mixed results in their impact on SETs (Surratt and Desselle 2007; Davidovitch and Dan Soen 2006; Okpala and Ellis 2005; Millea and Grimes 2002; Cashin 1995). Several researchers have attempted to distinguish how instructor's other non-verbal behavior impacts student perceptions of teaching quality (Babad, Avni-Baad and Rosenthal 2004 and Kierstead,, D;Agostino and Dill 1988).

Cashin (1995), reported that an instructor's personality was "not related to student ratings". Other researchers, however, have used a variety of personality traits of both the instructors and the student raters in an effort to explore and identify the most influential variables of personality (Helterbran 2008; Nerger et al. 2007; Clayson 2005; Clayson 1999; Kryzstofiak, Cardy, and Newman 1988). Students value, and thus give higher ratings to positive traits they define as caring for students, fairness related to grading, and accessibility (Hills, Naegle and Bartkus 2009; Surratt and Desselle 2007; Onwuegbuzie et al. 2007; Clayson and Sheffet 2006; Okpapa and Ellis 2005; Gursoy and Umbreit 2005; Jackson et al 1999; Hinkin 1991; Clayson and Haley 1990). Other valued personal qualities, include "exhibiting interest, if not passion, in one's teaching job, enthusiasm for teaching and the field of education, a sense of humor, and being approachable or "human" (Helterbran 2008). Students identify email with the instructor as a sign of social contact or a social relationship. Sheer and Fung (2007) found that "email communication affects teaching evaluations directly and feeds interpersonal relationships, which in turn, positively influence

teaching evaluations". Krzystofiak; Cardy and Newman (1988) studied many dimensions of college professor behaviors, showing that the relationship with students was paramount in impacting SETs.

Student perceptions have a direct impact on student attitudes when completing teaching evaluation forms. According to Curran and Rosen (2006), "course topic has just as strong an influence on attitude as does the instructor." Student attitudes regarding evaluations of teaching also depend on the student perceptions of the evaluation's use (El Hassan 2009; Brown 2008; Leite et al. 2006). Helterbran (2009) noted that students "have called for university generated formal student evaluation data to be put online for everyone to view." Students, desiring to express their perceptions of teaching quality for the use of other students, actively use the computer site Rate My Professor (www.ratemyprofessor.com) (Brown, Baillie and Fraser 2009; Felton, Mitchell and Stinson 2004).

According to Felton, Mitchell and Stinson (2004), "the sexier the instructor, the more difficult his or her class can be while obtaining high-marks on student evaluations." Riniolo et al. (2006), noted that "ratings for professors perceived as attractive rarely dropped below an average score (only 6 out of 211 scored below an average rating of 3 on a 5-point scale)." Carr, Davies and Lavin (2009) reported that, "the professional appearance and attire of the professor has a positive impact on the students' perceptions of a number of traits that are often considered in the evaluation of an academician."

Past researchers have shown that SETs are multidimensional and ratings are influenced by a number of external factors, beyond the control of any instructor (Johnson 2002; Read, Rama and Raghunandan 2001; Chasin 1995). There are three dimension clusters that appear consistently in SETs: "(a) instructor presentation of material, (b) facilitation of learning and (c) regulation of learning" (Jackson et al. 1999). Various other researchers have accepted and used these dimension clusters when developing or testing SETs to assist in producing the multidimensional aspect of the rating system (Onweugbuzie et al 2007; Marsh 2001). Researchers have identified various additional dimensions that impact SET results. Student mood (Munz and Fallert 1998) was correlated with both instructor and course ratings.

Our ongoing interest in student evaluations of teaching is their increasing importance in terms of faculty success and the lack of control that faculty members have over many of the factors that affect these evaluations. Our first research focused on identifying differences in student perceptions of distance education (DE) classes taught through videoconferencing. As we expected, remote-site students gave lower ratings to their DE classes than did originating-site students, and they were less satisfied with the classes in general (Uttley and Carson 2006). The goal of our next research was to identify external factors that influence SETs among faculty members teaching multiple sections of the same class. We concluded that if instructors wanted to maximize their evaluations, they would not teach more than two sections of a class in one semester. Counterintuitively, they would teach classes that meet early on Monday, Wednesday, and Friday mornings and they would never teach at 10:00 AM or 11:00 AM, two of our most popular class times. They would also teach upper level classes that are relatively large, and they would hope for a relatively low response rate on the SETs (Uttley and Carson 2007).

An additional element in the equation of SETs that is beyond the control of an instructor and only recently attracting attention is how seriously students take the evaluations when completing them. Onwuegbuzie et al. (2007) point out that the "lack of knowledge of the actual process that students use when they respond on TEFs (teaching evaluation forms) makes it difficult to claim that studies have provided sufficient evidence of substantive validity regarding TEF ratings" (p.118). Surratt and Desselle (2007) found that pharmacy students saw the SETs as an opportunity to express their opinions regarding a number of aspects of a course. Consistent with our interest in course factors beyond the control of instructors, and because of the lack of research regarding how students approach the teaching evaluation process, last year we began exploring two questions: "What are students' perceptions of teaching evaluations and do they take the evaluations seriously?" Given a couple of confusing questions (Carson and Uttley 2008) and additional literature, we are continuing the topic of student perceptions and attitudes this year. Further, this year we are exploring the impact on SETs of several additional variables:

(1) whether or not students like the instructor,

- (2) whether students feel strongly about the instructor,
- (3) whether students think the instructor teaches especially well or especially poorly,
- (4) whether students are influenced by the subject of classes, and
- (5) whether students think that the most effective instructors receive the best evaluations.

### METHODOLOGY

This research continues our examination of student perceptions and attitudes regarding student evaluations of teaching (SETs). In 2004, Lander University adopted the Individual Development and Educational Assessment (IDEA) instrument to collect student perceptions of all of their classes. The College of Business and Public Affairs (COBPA) began using this instrument in the fall of 1998. Unlike the other three colleges at Lander, our college has always used a unique strategy for conducting these student evaluations of teaching. On a particular two days, about 2/3 of the way through the semester, all classes in COBPA are evaluated. A faculty member, other than the class instructor, administers the evaluations at the beginning of class. The vast majority of the students in our college take four to six classes per semester, which means that by the end of the second day, a student has seen the same evaluation instrument for up to six times. As part of our continuing research into the impact of factors that are beyond an instructor's control on their Student Evaluations of Teaching (SETs), we asked students about their perceptions of the IDEA instrument and the process used by the College of Business and Public Affairs.

The IDEA instrument provides three summary measures for each class. The first measures effective teaching in terms of progress made on particular course objectives, which instructors pick from a list of 12 possible course objectives, designating each as essential, important, or not important to the particular course. Those objectives deemed essential are double weighted in the IDEA assessment calculations. Regardless of the objectives chosen by each instructor, the IDEA instrument is identical. Objectives are organized into six categories: *Basic Cognitive Background; Application of Learning; Expressiveness; Intellectual Development; Lifelong Learning;* and *Team Skills* (www.idea.ksu.edu). Most categories include multiple possible objectives, rated by students on a five-point scale from *no apparent progress* (1) to *exceptional progress* (5):

Basic Cognitive Background

- 1. Gaining factual knowledge (terminology, classifications, methods, trends)
- 2. Learning fundamental principles, generalizations, or theories

Application of Learning

- 3. Learning to *apply* course material (to improve thinking, problem solving, and decisions)
- 4. Developing specific skills, competencies, and points of view needed by professionals in the field most closely related to this course

Expressiveness

- 6. Developing creative capacities (writing, inventing, designing, performing in art, music, drama, etc.)
- 8. Developing skill in expressing oneself orally or in writing

Intellectual Development

- 7. Gaining a broader understanding and appreciation of intellectual/cultural activity (music, science, literature, etc.)
- 10. Developing a clearer understanding of, and commitment to, personal values
- 11. Learning to *analyze* and *critically evaluate* ideas, arguments, and points of view *Lifelong Learning* 
  - 9. Learning how to find and use resources for answering questions or solving problems
  - 12. Acquiring an interest in learning more by asking questions and seeking answers

#### Team Skills

5. Acquiring skills in working with others as a member of a team.

The second measure is based on the single statement, "Overall, I rate this instructor an excellent teacher." Response options range from *definitely false* (1) to *definitely true* (5). The third measure is based on the single statement, "Overall, I rate this course as excellent." The responses are the same as those used for the second measure. For each measure, the IDEA center calculates raw and adjusted average scores, recommending that the adjusted scores be used for comparisons. Scores are adjusted on the basis of students' professed desire to take the course, expressed effort put forth, and perceived amount of work required.

While the literature includes other studies of student perceptions of teaching evaluation instruments and student ideas about what should be included in evaluations of teaching, we designed our study specifically to focus on the evaluation instrument and process used at Lander. Within our college (COBPA) in the fall 2008 semester, faculty taught 82 classes from which to collect data. Since only COBPA employs the two-day evaluation process, we eliminated lower level general education courses (Anthropology 104, Sociology 101, Political Science 101 and 103, and Economics 101), which are taken by students from all colleges at Lander. We invited all college faculty to participate and distributed surveys to those who agreed. Approximately two-three weeks after the completion of the evaluation instruments, surveys were distributed to students. We asked that each student complete only one survey. Thus, as the distribution time lengthened, an increasing number of students from any specific class had already completed the survey. On the survey forms, we asked students to indicate whether they were at least 18 years old and whether we had permission to use their data in our analyses. A few students were evidently younger than 18 and a surprising number did not give permission for us to use their data. On last year's survey, these two questions were placed on a single horizontal line. We thought that students had just missed the second question giving us permission to use their survey because of the page layout. This year, we placed the question asking for permission to use the data below the question asking about age. We were confident that almost everyone would give permission to use their data. However, 62 students still did not give permission to use their data. Again, we eliminated those surveys, which left us with 353 usable surveys from students in 34 classes. Based on new literature, we added six new research questions. Our survey consisted of twelve perception questions and two demographic questions:

- 1. I believe the IDEA evaluation form is an accurate way for students to provide feedback to administrators about their professor's teaching objectives for the course.
- 2. I believe the IDEA evaluation form is an accurate way for students to provide feedback to professors regarding their teaching objectives for the course.
- 3. I believe the administration takes the IDEA evaluation results seriously for faculty retention, promotion, and salary increases.
- 4. I believe faculty take the results of the IDEA evaluations seriously.
- 5. I respond to the first IDEA evaluation on the first day in a serious manner; I read each question and carefully consider each of the responses, selecting the most appropriate.
- 6. I respond to the IDEA form on the second day in a serious manner. I read each question and carefully consider each of the responses, selecting the most appropriate. (reworded)
- 7. If I like a professor, their personality influences my responses in a positive manner. (new)
- 8. If I do not like a professor, their personality influences my responses in a negative manner. (new)
- 9. I am more likely to complete the evaluation when I really like or really dislike the professor (as opposed to having neutral feelings). (new)
- 10. I am more likely to complete the evaluation when I feel the professor taught especially well or especially poorly. (new)
- 11. The nature of the course (subject matter) influences my scoring on the evaluation. (new)

- 12. The professors who receive the best evaluations are not necessarily the most effective teachers. (new)
- 13.
   I am Female \_\_\_\_\_
   I am Male \_\_\_\_\_
- I am in the department of Business Administration \_\_\_\_\_
   I am in the department of Political and Social Sciences \_\_\_\_\_
   I am in another Department \_\_\_\_\_\_ (reworded)

The responses for each perception question were arranged in a Likert format from strongly disagree (1) to strongly agree (5). The demographic questions were checked or left blank to provide a yes or no format. We used the MicroCase<sup>\*</sup> statistical package to analyze the data. We compared students from the department of Business Administration to those from Political and Social Sciences (PaSS); we compared females to males; we compared stated seriousness from day one to day two. Since our data are nominal and ordinal level, we used contingency tables for our comparisons and rely on  $X^2$ , Somer's D and percentage differences for our analyses (Fox 2003).

## **DESCRIPTIVE ANALYSES**

Due to the lack of literature on student perceptions of the IDEA instrument, we have approached our exploration with a series of questions. Table 1 shows the demographic characteristics of our research population. Our survey percentage of females and males falls between the percentages for Lander and for our college. Respondents from both departments represent about 55% of their majors.

TABLE 1. Demographic characteristics of the research population.							
	Fem	nales	Males		Total		
Department Affiliation	Valid N	Row %	Valid N	Row %		% of Sample	
Business Administration	110	57.3	97	63.8	207	60.2	
Political and Social Sciences	55	63.8	44	40.4	99	28.8	
Neither/Both	27	14.1	11	12.5	38	11.0	
Total	192		152		344	100.0	

Table 2 includes the distribution of responses from our consenting respondents.

TABLE 2. Distribution of responses from the research population.						
	Strongly				Strongly	
Survey Questions	Disagree	Disagree	Neither	Agree	Agree	
Accurate way to provide feedback to						
administrators about professor's teaching	4.2	11.0	11.6	52	21.2	
objectives.						
Accurate way for students to provide feedback	15	12.1	0.0	52.8	20.6	
to professors about their teaching objectives.	4.5	12.1	9.9	52.8	20.0	
Believe administration takes evaluation						
results seriously for faculty retention,	5.6	15.5	27.1	35.9	15.8	
promotion, and salary increases.						
Believe faculty take the results of the IDEA	57	11 /	25.3	12.9	1/1 8	
evaluations seriously.	5.7	11.4	23.5	42.9	14.0	
Take first IDEA evaluation on the first day in a	18	5 1	137	37.3	30.0	
serious manner.	4.0	5.1	15.7	57.5	39.0	
Take last IDEA evaluation on the second day	4.0	97	15.4	37.0	33.0	
in a serious manner.	4.0	9.1	13.4	57.0	55.9	

<sup>&</sup>lt;sup>\*</sup> MicroCase Corporation, acquired in 1999 by Wadsworth, now a division of Thomson Learning, Inc.

TABLE 2. Distribution of responses from the research population, continued						
	Strongly				Strongly	
Survey Questions	Disagree	Disagree	Neither	Agree	Agree	
If like professor, their personality influences responses in a positive manner.	1.4	3.4	7.1	41.9	46.2	
If <b>do not</b> like a professor, their personality influences responses in a negative manner.	4.5	11.3	18.6	37.0	27.7	
More likely to complete the evaluation when really like or really dislike professor.	11.1	8.9	26.0	27.4	26.6	
More likely to complete the evaluation when feel professor taught especially well or especially poorly.	7.4	4.0	20.3	29.7	38.6	
The nature of the course (subject matter) influences scoring on the evaluation.	8.9	10.0	30.3	38.3	12.6	
The professors who receive the best evaluations are not necessarily the most effective teachers.	7.8	14.8	27.2	33.0	17.1	

#### **Replicated Questions:**

Consistent with our research from last year, findings indicate that over two thirds of our consenting respondents believe that the IDEA instrument is an accurate means to provide feedback about a course's teaching objectives to both administrators and professors. In both years, however, only about half of students thought that the administration was serious in its use of the assessment when determining retention, promotion and salary increases. About 60% of the queried students believe that the instructors took the results of the IDEA evaluation seriously. Table 3 shows the specific results. None of the differences are significant.

TABLE 3. Comparison of replicated questions.						
Survey Questions	2008 - % Agree	2009 - % Agree				
Accurate way to provide feedback to administrators about professor's teaching objectives.	70.8	73.2				
Accurate way for students to provide feedback to professors about their teaching objectives.	71.6	73.4				
Believe administration takes evaluation results seriously for faculty retention, promotion, and salary increases.	52.5	51.7				
Believe faculty take the results of the IDEA evaluations seriously.	62.5	57.7				

#### **Exploration Questions:**

For our comparative analyses, we collapsed the variable responses into three categories: disagree, neither disagree nor agree, and agree.

**Question 1:** Do students differ in how seriously they think about their completion of the IDEA instrument depending on whether it is the first day or the second day of the teaching evaluations process? Consistent with our results last year, it is obvious from figure 1 that over two thirds of all students say that they approach the evaluations in a serious manner. It is also obvious that fewer students take the evaluation seriously on the last day (70.9%) compared to the first day (76.3%). The relationship is significant based on Somer's  $D_{yx}$  of 0.75 (p<0.001), a proportional reduction in error (PRE) measure, that indicates knowing a student's seriousness on day one reduces the error in predicting their seriousness on day two by 75%. Six percent of students, who stated that they approached the evaluations seriously on

the first day, said that they did not approach them as seriously on the second day, the change we expected to find.



# FIGURE 1. Distribution of students showing the seriousness with which they approach teaching evaluations for the first day and second day of the evaluation period.

**Question 2:** Do students differ in how seriously they think about their completion of the IDEA instrument based on whether they are majors in the department of business administration or political and social sciences (PaSS)? Kwan (1999) and Nerger et al. (2007) found academic discipline to be associated with differences in SETs. Do these differences reflect different perceptions about the teaching evaluation process or instrument as well? Figure 2 shows the comparison of students from business and PaSS. While a higher percentage of PaSS students (79%) than business students (77.7%) state that they approach the teaching evaluation process in a serious manner, the difference is not significant, a finding which is consistent with last year's research.



# FIGURE 2. Comparison of students from business and PaSS showing the seriousness with which they approach teaching evaluations for the first day of the evaluation period.

**Question 3:** Do females differ from males in how seriously they approach their completion of the IDEA *instrument*? In general, females tend to evaluate teaching more positively (Millea and Grimes 2002; Davidovitch and Dan Soen 2006). In figure 3, it is easy to see that significantly (p=0.026) more females (82.3%) approach the evaluation process seriously than do males (70.1%) on the first day of evaluations. We found no other significant differences between female and male students.



FIGURE 3. Comparison of females and males showing the seriousness with which they approach teaching evaluations for the first day of the evaluation period.

**Question 4:** Are student evaluations influenced by whether they like or do not like a professor? The inconsistent findings reported in the literature on the affects of personality on SETs, prompted us to ask two questions: one, does liking a professor positively influence a student rating; and two, does not liking a professor negatively influence a student rating. Nearly 9 out of every 10 students (88%) agreed that they are more likely to rate a professor positively if they like that professor. This result is far higher than the 65% of all students who agreed that not liking a professor influences them to give a lower rating. With respect to the influence of not liking a professor, however, students from the two departments differ significantly (chi-square = 12.056, p<0.001) as seen in figure 4. Among business students, 70%, compared to 53% of PaSS students agreed that they tended to give negative ratings to professors they did not like. Departmental differences between both females (p=0.042) and males (p=0.027) are also significant. In both departments, women are more likely than men to give negative ratings when they do not like a professor. In all of our explorations, we found no other significant differences between the females from the two departments.



FIGURE 4. Comparison of students from business and PaSS showing the influence that not liking a professor has on the student ratings.

**Question 5:** Are students more likely to complete an evaluation when they feel strongly (either liking or disliking) a professor? This question was related to the last two questions. While 55.6% of Business students compared to 48% of PaSS students agreed that strong feelings for a professor made them more likely to complete an evaluation, the difference among students was not significant.

**Question 6:** Are students more likely to complete an evaluation when they believe that a professor teaches especially well or especially poorly? This question was related to the last three questions. Among all students, 68.3% agreed that they were more likely to complete an evaluation if they felt strongly about a professor's teaching ability. Department differences among students were not significant.

**Question 7:** Are student evaluations influenced by the subject matter of the course? Overall, 50.9% of students agreed that topic of a course influenced their evaluations. However, department differences were significant (p=0.01, based on chi-square = 6.188) with 44.9% of Business students compared to 59% of PaSS students agreeing that subject matter influenced their evaluations.



FIGURE 5. Comparison of students from business and PaSS showing the influence that subject matter has on the student ratings.

**Question 8:** Do students think that the most effective instructors receive the best evaluations? Overall, 50.9% of students believed that the most effective teachers do not necessarily receive the best evaluations. Department differences were not significant.

## DISCUSSION

The goal of this study was to determine what impact, if any, student perceptions have on the evaluation process of the teaching quality. A short survey using Likert-scale questions was developed to assess students' seriousness and perceptions regarding the use of the student evaluation of teaching results. This survey was administered to students approximately one week after they completed the formal student evaluation of teaching.

We believe that we fixed the problems associated with our survey instrument from last year when we were surprised by the number of students that did not give us permission to use their survey responses in our research. We thought that the layout of the survey instrument had led students to miss this question. However, after redesigning the layout, 15% of students definitely denied us permission to use their data. Since students have no particular incentive to complete our surveys, we do not understand why students would take time to complete the questionnaire, but then deny us permission to use their data.

Last year, our second set of identifiers was also problematic. We asked students to indicate their academic affiliation. A few students checked that they were in either the department of business administration or political and social sciences and also checked that they were not in our college. We did not realize that students did not know the name of the college in which their department was housed. This year, we revised that question and asked students if they were in a department other than business administration or political and social sciences, rather than asking if they were in a different college.

About 11% of students stated that they were in either both departments or were in neither department. Students could be in both departments if they have a major in one and a minor in the other or if they are taking electives in the non-major department. Almost no students at Lander double major. Several specific classes in sociology are quite popular as electives among students throughout the university.

Last year while entering data, we noticed a contradiction in a sequence of answers. About 10% of students, who said that they did not take the evaluations seriously on the first day, stated that they approached the second-day evaluations with equal seriousness. This problem occurred because of the wording of our questions. This year we reworded both of these questions to simply ask how seriously students approached the evaluations on each day. The answers could have ranged from very seriously to not at all seriously in a Likert format. Even with this year's better-worded question about second day evaluations, 5.5% fewer students stated that they gave the same serious attention to the assessments completed on day two as to those completed on day one of the evaluation period. We believe this occurs from a fatigue factor given that students saw the same instrument five to six times during that two-day period. This tends to introduce a systematic bias, especially for those instructors who teach primarily on Tuesdays and Thursdays since the student evaluations have always been done on Mondays and Tuesdays. Our experiences as proctors for the student evaluations lead us to question the stated seriousness with which students say they approach the evaluations. We observe some students marking designs with their responses or marking all responses in a single column of the form. We hear statements from students such as, "I really like you as a teacher so I gave you all 1s", (1 is strongly disagree).

Last year, we did not include an open-ended question asking for comments, an oversight. We rectified that oversight this year. However, we received only a few comments, which were inconsequential. Nonetheless, future questionnaires will also include a place for students to add comments.

Our most recent findings again indicate that slightly better than 70% of our participating students believe that the IDEA instrument is an accurate means of providing feedback about class teaching objectives to the administration as well as to the professor. Conversely, only slightly more than 50% of the students perceive that the administration is serious in its use of the evaluation when determining retention, promotion and salary increases. This year, about 58%, of the queried students believe that the instructors took the results of the IDEA evaluation seriously compared to about 71.5% last year.

Several studies, Nerger et al. 2007; Onwuegbuzie et al. 2007; Clayson and Sheffet 2006; Okpapa and Ellis 2005; Clayson 1999; Clayson and Haley 1990, identified that characteristics of an instructor's personality are directly correlated to both instructor and course ratings. In the same area of reference, Surrey and Desselle (2007) found that students were more willing to complete the teaching evaluation if they really liked or disliked the professor. This year, our research focused on students' opinions of their instructors and how those opinions influenced their evaluations. Nearly 90% of students agreed that they are more likely to rate a professor positively if they like the professor. This result is far higher than the 65% of students who agreed that not liking a professor influences them to give a lower rating. Students from the two departments differ significantly in their responses about the influence of not liking their instructors. Only about half of PaSS students compared to over two thirds of business students agreed that they tended to give negative ratings to professors they did not like. Differences between females and males from the two departments are also significant. Among PaSS students, both females and males are less likely than females and males from business to be influenced negatively by not liking a professor. We expected the relationship between not liking a professor and writing a negative evaluation to be stronger than the relationship between writing a positive evaluation and liking a professor and were a bit surprised by this result. While students' feelings about a professor influenced whether their evaluations were positive or negative, they did not influence whether or not students completed evaluations. We think this is because these evaluations occur at the beginning of each class period, with a faculty proctor present. Certainly, a student could put the evaluation down on their desk and work on other class work, but it would be obvious to everyone in the room. We suspect that liking of not liking an instructor would have a greater effect on completion of the evaluations at schools where students complete the evaluations online.

Related to our questions about liking or not liking a professor, are questions about student perceptions about professors who teach well or poorly. We asked if students were more likely to complete an evaluation when they believe that a professor teaches especially well or especially poorly. About two thirds agreed that they were more likely to complete an evaluation when they felt strongly about a professor's teaching ability. This is a result that we expected. We hope, however, that students use the evaluations to do more than affirm their beliefs about teaching competency. All student evaluations of teaching give students a chance to provide feedback about teaching: what worked well, what was confusing, what did not work, and suggestions to make the class better. We also asked if students think that the most effective instructors receive the best evaluations, a desirable goal if SETs are valid. Unfortunately, slightly under half of the students believed that this was the outcome, a disappointing result, and one that bears future exploration and makes relying on the SETS questionable.

Our last question, based on the literature, asked if student evaluations are influenced by the subject matter of the course. We were not surprised to see students influenced by the topic of a class when they complete their evaluations. The surprise was the significant (15%) difference between business students compared to Political and Social Science students.

Our research over the last several years continues to raise questions about student evaluations of teaching in general. The increasing pressure on universities to treat students as customers suggests that ethically, potential students and their parents should have access to these evaluations of teaching in the name of full disclosure, just as they get information on crimes on campus. However, given continuing questions about the validity of student evaluations, potential customers would also need information on factors that influence SETs. How would all that be presented and how would potential students and their parents make sense of it all? We believe that this research continues to challenge the received wisdom regarding the validity and reliability of the IDEA evaluation instrument. The need to implement the most valid and reliable instrument to gauge teaching quality is vital in the world of accountability.

#### REFERENCES

- [1] Aultman, L. 2006. An unexpected benefit of formative student evaluations. (QUICK FIX). *College Teaching* 54.3:251.
- [2] Addison, W.E., J. Best, and J.D. Warrington, 2006. Student's perceptions of course difficulty and their ratings of the instructor. *College Student Journal* 40(2):409-416.
- [3] Anderson, K. J. and G. Smith . 2005 Students' preconceptions of professors: Benefits and barriers according to ethnicity and gender. *Hispanic Journal of Behavioral Science* 27(2):184-201.
- [4] Babad E., D. Avni-Baad; and R. Rosenthal. 2004. Prediction of student's evaluations from brief instances of professors' nonverbal behavior in defined instructional situations. *Social Psychology of Education* 7:3-33.
- [5] Basow, S.A. 1995. Student evaluations of college professors: When gender matters. *Journal of Educational Psychology* 87(4):656-665.
- [6] Birnbaum, M.1999. A survey of faculty opinions concerning student evaluations of teaching. <u>The Senate Forum: A publication of the Academic Senate of California State University, Fullerton</u>, <u>14</u>:19-22.
- [7] Brown, M. J. 2008. Student perceptions of teaching evaluations. *Journal of Instructional Psychology* 35(2):177-181.
- [8] Brown, M.J., M. Baillie and S. Fraser. 2009. Rating ratemyprofessors.com. A comparison of online and official student evaluations of teaching. *College Teaching* 57(2):89-92.
- [9] Burden, P. 2009. A case study into teacher perceptions of the introduction of student evaluation of teaching surveys (SETs) in Japanese tertiary education. *Asian EFL Journal 11*(1):126-149.
- [10] Burdsal, C. A. and J. W. Bardo. 1986. Measuring Student's Perceptions of Teaching: Dimensions of Evaluation. *Educational and Psychological Measurement*. 46(1):63-79.
- [11] Carr, D., T. Davies; and A. Lavin. 2009. The effect of business faculty attire on student perceptions the quality of instruction and program quality. *College Student Journal 43*(1):45-55.
- [12] Carson, L.A. and M. Uttley. 2008. Student Perceptions of Teaching Evaluations. Southeast InfORMS annual conference, 44, *Southeast InfORMS Proceedings*, 44:256-267.
- [13] Cashin, W. E. 1989. Defining and Evaluating College Teaching. IDEA paper, No. 21, September. Center for Faculty Evaluation & Development. Kansas State University
- [14] Cashin, W. E. 1995. Defining and Evaluating College Teaching. IDEA paper, No. 32, September. Center for Faculty Evaluation & Development. Kansas State University
- [15] Cashin, W., and R. Downey. 1992. Using global student rating items for summative evaluation. *Journal of Educational Psychology* 84:563-572.
- [16] Clayson, D.E. 2005. Within-class variability in student-teacher evaluations: Examples and problems. *Decision Sciences Journal of Innovative Education* 3(1):109-124.
- [17] Clayson, D. E. and D. Haley 1990. Student evaluations in marketing: What is actually being measured? *Journal of Marketing Education* 12(9):9-17.
- [18] Clayson, D. E. and M. J. Sheffet, 2006. Personality and the Student Evaluation of Teaching. *Journal of Marketing Education* 28(2):149-160.
- [19] Curran, J. M. and D. E. Rosen. 2006. Student attitudes toward college courses: An examination of influences and intentions. *Journal of Marketing Education* 28(2):135-148.
- [20] Davidovitch, N. and D Soen. 2006. Class attendance and students' evaluation of their college instructors. *College Student Journal* 40.3:691(13).
- [21] El Hassan, K. 2009. Investigating substantive and consequential validity of student ratings of instruction. *Higher Education Research & Development* 28(3):319-333.
- [22] Felton, J., J. Mitchell; and M. Stinson. 2004. Web-based student evaluations of professors: the relations between perceived quality, easiness and sexiness. Assessment & Evaluation in Higher Education 29(1):91-108.
- [23] Fernandez, J., M. Mateo &, J. Muniz 1998. Is there a Relationship between Class Size and Student Ratings of Teaching Quality? *Educational and Psychological Measurement* 58(4):596-604.

- [24] Fernandez, J. and M. A. Mateo. 1992. Student evaluation of university teaching quality: Analysis of a questionnaire for a sample of university students in Spain. *Educational and Psychological Measurement* 52(3):675-686.
- [25] Fox, William. 2003. Social Statistics, fourth edition. Belmont, California: Wadsworth Group.
- [26] Germain, M. and T. Scandura. 2005. Grade inflation and student individual differences as systematic bias in faculty evaluations. *Journal of Instructional Psychology* 32.1:58(10).
- [27] Grayson, J. P. 2004. The relationship between grades and academic program satisfaction over four years of study. *The Canadian Journal of Higher Education* 34(2):1-34.
- [28] Gursoy, D. and W.T. Umbreit. 2005. Exploring Students' Evaluations of Teaching Effectiveness: What Factors are Important? *Journal of Hospitality & Tourism Research*. 29(1):91-109.
- [29] Haskell, R. 1997. Academic Freedom, Promotion, Reappointment, Tenure and The Administrative Use of Student Evaluation of Faculty: (Part II) Views From the Court. *Education Policy Analysis Archives 5*(17).
- [30] Helterbran, V. R. 2008. The ideal professor: Student perceptions of effective instructor practices, attitudes, and skills. *Education* 129(1):126-138.
- [31] Hills, S. B.; N. Naegle and K. R. Bartkus. 2009. How important are items on a student evaluation? A study of item salience. *Journal of Education for Business* 84(5):297-303.
- [32] Hinkin, T. R. 1991. The Effects of Time of Day on Student Teaching Evaluations: Perception versus Reality. *Journal of Management Education* 15(1):35-45.
- [33] Horner, K. L., H. G. Murray and J. P. Rushton. 1989. Relation between aging and rated teaching effectiveness of academic psychologists. *Psychology and Aging* 4(2): 226-229.
- [34] Hoyt, D. and S. Perera. 2000. Validity of the IDEA Student Ratings of Instruction System: An Update. *IDEA Research Report #2*. Manhattan, KS: Kansas State University, Center for Faculty Evaluation and Development.
- [35] Jackson, D.; C. Teal; S. Raines; T. Nansel; R. Force, and C. Burdsal. 1999. The Dimensions of Students' Perceptions of Teaching Effectiveness. *Educational and Psychological Measurement* 59:580-596.
- [36] Johnson, V. 2002. Teacher Course Evaluations and Student Grades: An Academic Tango. *Chance* 15:9-16.
- [37] Kierstead, D., P. D'Agostino and H. Dill. 1988. Sex role stereotyping of college professors: Bias in students' rating of instructors. *Journal of Educational Psychology* 80(3):342-344.
- [38] Krzystofiak, F., R. Cardy and J. Newman. 1988. Implicit personality and performance appraisal: The influence of trait inferences on evaluations of behavior. *Journal of Applied Psychology* 73(3):515-521.
- [39] Kwan, K. 1999. How fair are student ratings in assessing the teaching performance of university teachers? *Assessment of Evaluation in Higher Education* 24:181-195.
- [40] Leite, D., R. A. Santiago, C. S. Sarrico, C. L. Leite and M. Polidori. 2006. Students' perceptions on the influence of institutional evaluation on universities. Assessment & Evaluation in Higher Education 31(6):625-638.
- [41] Marsh, H. W. 2001. Distinguishing Between Good (Useful) and Bad Workloads on Students' Evaluations of Teaching. *American Educational Research Journal 38*(1):183-212.
- [42] Maurer, T., J. Beasley, J. Dilworth, A. Hall, J. Kropp, M. Rouse-Arnett, and J. Taulbee. 2006. Child and family development students polled: study examines student course evaluations. *Journal of Family and Consumer Sciences* 98.2:39-45.
- [43] Millea, M., and E. Grimes. 2002. Grade expectations and student evaluation of teaching. *College Student Journal* 36:582-590.
- [44] Nerger, J.; W. Viney, and R. Riedel II. 1997. Student ratings of teacher effectiveness: use and misuse. *The Midwest Quarterly* 38:218(16).
- [45] Okpala, C. O. and R. Ellis. 2005. The perceptions of college students on teacher quality: A focus on teacher qualifications. *Education* 126(2)L 374-383.

- [46] Onwuegbuzie, A. J.; A. E.Witcher; K. M. T.Collins; J. D.Filer; C. D. Wiedmaier, and C. W. Moor. 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.
- [47] Pike, C. 1998. A validation study of an instrument designed to measure teaching effectiveness. *Journal of Social Work Education 34*:261-271.
- [48] Read, W.J; D.V. Rama; K. Raghunandan. 2001. The relationship between student evaluations of teaching and faculty evaluations. Journal of Education for Business *76*(4):189-192.
- [49] Rifkin, T. 1995. Eric Review: Faculty Evaluation in Community Colleges. *Community College Review* 01061995:63-72.
- [50] Riniolo, T. C., K. C. Johnson, T. R. Sherman and J. A. Misso. 2006. Hot or not: Do professors perceived as physically attractive receive higher student evaluations? *The Journal of General Psychology* 133(1):19-35.
- [51] Schneider, B., P. J. Hanges, H. W. Goldstein and E. P. Braverman. 1994. Do customer service perceptions generalize? The case of student and chair ratings of faculty effectiveness. *Journal of Applied Psychology* 79(5):685-690.
- [52] Schulze, E. and A. Tomal. 2006. The chilly classroom. Beyond gender. *College Teaching* 54(3):263-269.
- [53] Seldin, P. 1999. Current practices—good and bad—nationally. In P. Seldin (Ed.), *Changing Practices in Evaluating Teaching* (pp.1-24). Bolton, MA: Anker Publishing Company.
- [54] Seldin, P. 1993. The Use and Abuse of Student Ratings of Professors. *The Chronicle of Higher Education* 39:40-42, July 21.
- [55] Sheer, V. C. and T. K. Fung. 2007. Can email communication enhance professor-student relationship and student evaluation of professor?; Some empirical evidence. J. Educational Computing Research 37(3):289-306.
- [56] Sprinkle, J. E. 2008. Student perceptions of effectiveness: An examination of the influence of student biases. *College Student Journal* 42(2):276-293.
- [57] Surratt, C.K. and S.P. Desselle. 2007. Pharmacy student's perceptions of a teaching evaluation process. *American Journal of Pharmaceutical Education* 71(1):1-7.
- [58] UCSB. 2004. Teaching large classes. Office of Academic Programs, Instructional Development University of California, Santa Barbara. <u>http://www.oic.id.ucsb.edu/Resources/Teaching/Large.ucsb.html</u>
- [59] Uttley, M. and L.A. Carson. 2007. Student evaluations of teaching: Multiple class sections compared. Southeast InfORMS annual conference,43, *Southeast InfORMS Proceedings*, 43:264-277.
- [60] Uttley M. and L.A Carson. 2006. Student Perceptions of Classes Taught by Videoconference: Local and remote sites compared. Southeast InfORMS annual conference, 42, *Southeast InfORMS Proceedings*, 42:194-205.