THE INTERNATIONAL JOURNAL OF HUMANITIES & SOCIAL STUDIES

Factors Determining Lecturers Perception of Student Evaluation of Teaching in Kenyan Universities

Lucy W. Gichinga
Jomo Kenyatta University of Agriculture and Technology, Kenya
Elegwa Mukulu
Professor, Jomo Kenyatta University of Agriculture and Technology, Kenya
Dr. Samuel Mwachiro
Pwani University, Kenya

Abstract:
Student evaluation of teaching (SET) have become a common method of lecturer evaluation as it offers a clear perspective of a lecturers effectiveness and is an important tool for lecturers who wish to improve their performance. However when used without other supportive tools, students become the sole determinants of the success or failure of a lecturer despite lecturers viewing the student evaluation of teaching as lacking in validity and reliability. This study seeks to establish factors that determine the lecturers' perception of the student evaluation of a lecturers teaching in Kenyan universities. The study objectives were to assess the extent to which lecturers' perception was influenced by gender, situational factors and purpose/use of the evaluation. The methodology included survey or descriptive design. Universities were stratified into public and private. The target population consisted of full time lecturers numbering 1114 drawn from two public and two private universities selected using stratified purposive sampling. Respondents from various schools were then sampled using simple random method. Data was collected using semi-structured questionnaires and analyzed using statistical analysis generated using the computer application package SPSS version 11.5. Several methods used to analyze data included descriptive statistics, cross tabulation while hypotheses were tested using t-test and f-test approach. The findings indicated that gender did greatly influence perception to student based evaluation where female academics were found to be more sensitive to gender related constructs while male academics were less sensitive. Most academic staff was dissatisfied with the use of SETs for summative purposes but was more agreeable to use of SETs for formative purposes. Majority of respondents felt that SETs were not fair valid and reliable although generally situational factors did not have a strong effect of perception of lecturers towards student based evaluation. The study recommends that organizations design and use performance evaluation methods that ensure equal opportunities for men and women to narrow the gender based bias. It is wise for organizations to blend SETs along with other evaluation methods to increase their validity and reliability in order to gain more acceptances. Evaluations could be done mid semester to enable academics make changes to improve delivery of courses in which they are being evaluated.

Keywords: *Student Evaluation of Teaching, Gender, Situational Factors, Formative and Summative Uses of Evaluation*

1. Introduction

Performance evaluation assesses the performance against pre-determined measures of performance, based on key success factors which may include measures of deviation from the norm, tracking past achievements and measures of output and input (Millmore, et al. 2007). De Nisi and Gonzales (2000) concur that a central goal of performance evaluation is to increase performance at the individual and, subsequently, the organizational level. The dilemma for universities is whether performance appraisal systems can channel the efforts of employees in an organizationally relevant way while recognizing staff concerns for continuing professional development and academic freedom. European universities for a long time had the belief that universities were autonomous, liberal academies committed to independence, neutrality and the advancement of knowledge. Shahzad et al. (2008) suggest that employee commitment and productivity can be greatly improved with performance evaluation. Universities are now being subjected to ever increasing levels of accountability, part of which has involved the widespread application of performance evaluation systems. Students are now viewed as 'clients', deans as 'managers' and knowledge marketed as a commodity.

According to Iyamu and Aduwa (2005) SETs are periodic evaluation of teachers performance which involves systematic gathering and analysis of information on the basis of which decisions are taken regarding the effectiveness, efficiency and/or competence of the lecturer in realizing set professional goals and the desire of the university to promote effective learning. Informal student evaluation of teaching (SET) began in Italy and lecturers were paid according to teaching abilities (Barette et al. 2006). According to Marsh and Bailey (1993) students were asked to describe lecturers according to what they considered

effective and also characteristics of good teachers and scales were used based on qualities believed to be important for teaching (Barette 2006). SETs have now become a common method of lecturer evaluation as it offers an indispensable perspective of a lecturer's effectiveness and an important tool for lecturers who wish to improve their performance (Stoklasa et al. 2011). Students are one of the consumer groups interested in the product of a college education and their opinions are considered a vital source of information concerning the quality of instruction at colleges and universities. Moreover students have demanded greater transparency around the outcomes of evaluation and teachers response to them.

Emery et al. (2003) opines that student evaluation of teaching effectiveness has been on the increase in colleges and universities in the USA and is often the most influential information in promotion and tenure decision at colleges and universities focused on teaching. The common methods for SETs are measurement tests, teacher characteristics, student achievement test scores, classroom performance (Imogie 2000) and according to Gold (2001) student evaluation of teachers have been largely unorganized. Areas commonly evaluated are organization and planning, lecturer student rapport, clarity and communication, grading of students, flexibility of approaches towards teaching and completion of course content (Stoklasa et al. 2011).

SETs are important for improving efficiency and equity which depends on teacher skill, resources and motivation (Weinberg 2007). Continuous improvement is possible through effective monitoring and evaluation as it ensures teachers' strengths are known and gaps addressed in order to raise standards according to Santiago and Benavides (2009). Iyamu and Aduwa (2005) add that SETs are important as they increase chances of excellence in teaching, provides means of participation between students and teachers in the teaching-learning process, provides the only direct extensive information about faculty members teaching and acts as an indication that faculty and administration recognize the importance of student involvement in shaping the institutions goals and practices. Students are better able to judge classroom atmosphere, pace of instruction and clarity and organization of class material (Belanger 2009).

In Malaysia university lecturers' performance assessment is carried out every semester for all courses and sections within courses offered for both undergraduate and postgraduate programs (Samian and Noor 2012). The main purpose of the evaluation is to provide information and feedback to academic staff on the teaching performance which in turn should enable them to do some self reflection and eventually take necessary actions to enhance their teaching performance in the future.

Samian and Noor (2012) add that evaluation takes place several weeks towards the end of the semester to evaluate delivery of courses taken. Data collected from the students' responses is analyzed, tabulated and presented to university top management and based on the finding remedial actions and future plans are strategized to ensure continuous improvement in the quality of teaching and learning. Penny (2003) adds that the evaluation reports are used for promotional exercises and strategic decisions such as retention, tenure, curriculum development, external quality care and research on teaching. Lecturers are also able to log into the system to see their own results and students comments at the end of the semester which can be used for self reflection and improvement. Yosef et al. (2012) found that student evaluation is the most widely used instrument to measure lecturers' performance.

Universities in Kenya have developed variants of performance evaluation systems for use in respective institutions. Ngware and Ndirangu (2007) identified student based appraisal as the most widely used in the universities even though other techniques such as peer reviews are also used. The SETs are done towards the end of the semester and results analyzed for use in deciding course loading for the subsequent semester. However lecturers do not get to know the results of the evaluation as there is no discussion between them and the administration. Kenya has 22 chartered public universities, and 9 constituent colleges, 17 chartered private universities and 5 private constituent colleges (CHE 2014). The focus of the study was two public and two private universities. In each category one university was relatively old while the other was relatively young. The study targeted full time academic/teaching staff in these universities as it sought to explore factors that determine perception of student evaluation of lecturers.

1.1. Statement of the Problem

Performance evaluation systems used in universities have largely relied on student based academic staff appraisal carried out towards the end of the semester and analyzed per semester to determine course loading for the subsequent semester. Student evaluation of lecturers teaching in universities has been used over the years to measure lecturer effectiveness and has been as an effective tool to measure effectiveness of quality (Spooren and Mortelmans 2006). Newly employed academic staff begins to teach without being attached to a senior lecturer for mentoring and student evaluation of their performance becomes useful as students are able to provide the department a lot of information about teaching by their lecturer than the head of department or faculty dean.

However when used in isolation and without other supportive tools, students become the sole determinants of the success or failure of a lecturer (Seldin 1993). Lecturers across universities and colleges view the student evaluation of teaching as lacking in validity and reliability as they consider students too immature to evaluate quality of teaching due to limited subject knowledge which may affect judgment making. Besides enough time must pass before students can be in a position to effectively evaluate effectiveness of the teaching and learning experiences as the evaluation is done in the middle of the semester. The student evaluation tool may contain irrelevant variables which may interfere with students' perception. Moreover lecturers do not get to know the results of the evaluation as there is no discussion between them and the administration.

This study seeks to establish factors that determine the lecturer's perception of the student evaluation of a lecturers teaching. However students may not be in a position to discern the quality or validity of the lecture content as they are usually influenced more by the style of delivery than by the quality of the content. The evaluation report is given directly to the head of department minimizing lecturers' participation in their own appraisal and reducing their intrinsic motivation which would facilitate growth and development (Ngware and Ndirangu 2007). According to Ng'ang'a (2012) Kenyan universities have slipped in ranking

worldwide indicating a low level of lecturer performance and consequently low competitive advantage. Poor appraisal systems have led to significant capacity problems in some faculties and affected the student-lecturer ratio.

1.2. Study Objectives

1.2.1. General Objective

The general objective of this study is to establish factors that influence lecturers' perception of student evaluation of teaching (SETs).

1.2.2. Specific Objectives

- To assess the extent to which gender influences perception of student evaluation of teaching SETs.
- To assess the extent to which purpose/use of the evaluation results influences lecturers' perception of student evaluation of teaching SETs.
- To examine the extent to which situational factors influences lecturers' perception of student evaluation of teaching.

1.3. Significance of Study

This study will benefit students who are direct beneficiaries of teaching, lecturers and university administrators as they endeavour to improve service delivery to the students, improvement of higher education and also the role of the university in achieving Vision 2030. The study will also stimulate further discussions in the academic circles regarding SETs against other evaluation approaches.

2. Literature Review

2.1. Theoretical Framework

Performance evaluation in private and public sector organizations has been studied from a variety of perspectives. The theoretical framework of this study is based on the equity theory.

2.1.1. Equity Theory

This theory posits that individuals compare the ratio of their own outcomes and inputs to the ratios of outcomes and inputs of referent others. If individuals perceive these ratios as unequal, they will be motivated to take actions designed to restore equity. The key components of equity theory are contributions that employees bring to a job situation, rewards or punishments that employees receive from a job situation; the individuals with whom employees compare their ratio of outcomes to inputs, the comparisons that employees make that determine whether they perceive themselves to be in an equitable situation or an inequitable situation; and the actions that individuals take in an attempt to restore equity (Adams 1963, 1965 and Mowday 1991).

Individual Outcomes	vs	Comparison Person Outcomes
Individual Inputs		Comparison Person Inputs

Figure 1: Equity Theory

Employees seek to maintain equity between perceived inputs that they contribute to a job and the perceived outcomes that they receive from the organization against perceived inputs and outcomes of others. A person will feel demotivated if they feel unfair treatment compared to others inside and outside the organization that employs him. An employee perceives equity if he perceives the ratio of his inputs to his outcomes to be equal to those inside and outside the organization. The employee will accept it if a comparable employee receives more input if this comparable employee contributes more input. Employee will always seek fair treatment but if they perceive unfairness, they are likely to feel distressed and humiliated.

According to Boyd and Kyle (2004) distributive justice refers to the fairness of reward in the light of an employee's performance while procedural justice to the accuracy and suitability of evaluation processes and procedures used to determine outcomes. A study by Newman and Milkovich (1990) showed that there are considerable gaps in procedural justice, especially in terms of measuring external market wages for determining external job value, although modern compensation systems tend to offer reinforced safeguards that should offer gradual increase in procedural justice and more justified connections between the labor market situation and employees' compensation. It has also been the subject of some criticism in the sense that it is underspecified , it is better able to predict reactions to situations in which individuals are under-rewarded in comparison to situations in which they are over-rewarded, it is based on an equity rule in that those who contribute the most should receive the highest outcomes rather than other types of rules that might determine resource allocations, such as equality, needs or hierarchy and it focuses on distributive justice and fails to consider the importance of procedural justice or interactional justice (Ambrose & Kulik, 1999). Performance evaluation should provide employees a voice in the process such that if employees are confident in the fairness of the process they are likely to accept the outcome of a judgment even when it does not favour them (Gary 2003).

2.2. Conceptual Framework

The conceptual framework will cover the following independent variables; gender, purpose of student evaluation and situational factors that influence the lecturers' perception of student evaluation of teaching in Kenyan universities.

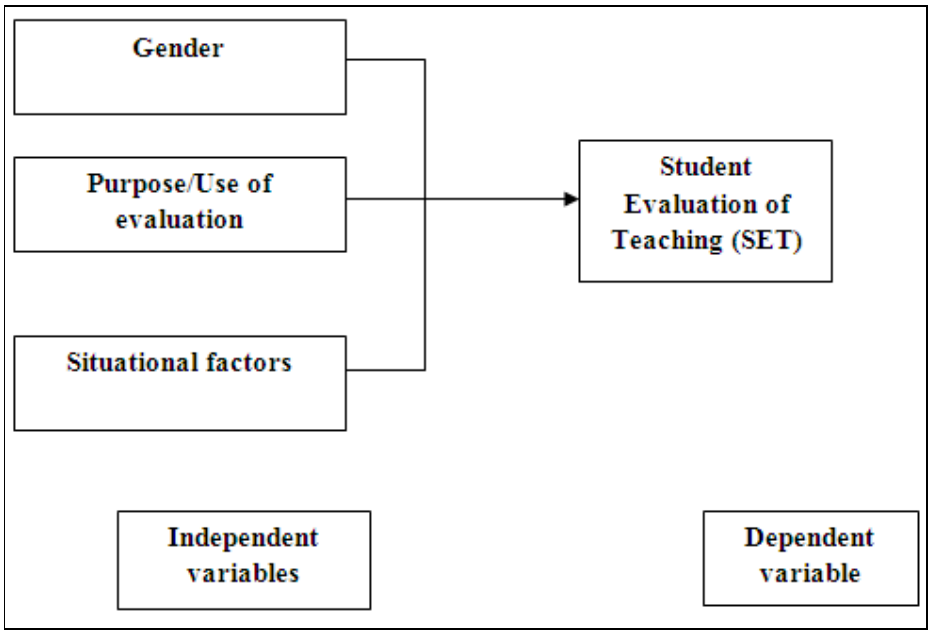

Figure 2: Conceptual Framework

2.2.1. Influence of Gender on Lecturer' Perception of Student Evaluation of Teaching

According to studies carried out in Malaysia, gender was found to have a significant effect on lecturers' perception of students' evaluation Harun, et al. (2011). Benz and Blatt (1996) felt that student judgment about women academics capability seem persistently to be contaminated by judgment of their womanliness to a point where insignificant issues such as audibility are subject to stereotyping. According to Carson (2001) majority of female lecturers considered student –lecturer evaluation to be gender biased against female lecturers. Female lecturers felt they have to work harder to prove their academic authority than their male counterparts and students rated attractive female lecturers more favourably (Bagilhole and Woodward 1995). According to Heckert et al. (1999) lecturers reported that student behaviour was indicative of female faculty members experiencing relative lack of respect. The female lecturers also felt that students systematically misattributed academic qualifications of male and female faculty to the detriment of female staff (Miller and Chamberlain 2000). Barnes and Letherby (1998) found that women academics faced greater demands to compete as researchers and nurture as women while Carson (2001) adds that female lecturers felt pressure from students and male colleagues to play a stereotypically feminine role. Yoder and Schleicher (1996) also found that women working in gender incongruent occupations were judged by students as competent. Women identified student prejudice as a largely male phenomenon with male students more likely to openly challenge their authority than female students (Bagihole 1995).

2.2.2. Purposes of Evaluation

In his findings Kelley (2012) suggests that student evaluation is widely used at most universities and colleges for the purpose of providing feedback to faculty in order to help them improve the teaching or alter course content (Formative purpose); used in merit and promotional decisions, course selection for students administrative purpose. Gravestock and Greenleaf (2008) add that summative purposes are widely used in many institutions. Virtually all colleges use student evaluation of instructors as a measure of instructor performance (Magner 1997), therefore, such student evaluations have a significant impact on tenure, promotion or merit pay decisions concerning faculty (Ehie and Karathanos, 1994). Feedback from students may help instructors to improve their teaching performance (Marsh 1991). Unfortunately, the use of such ratings for evaluations relating to reward systems of a college or university may be problematic.

2.2.3. Situational Factors

Validity is the degree to which a measure measures what it is supposed to measure. The major aspect of validity in performance evaluation is content validity. An appraisal instrument has content validity to the extent that it includes most of the important job behaviors and/or results of the job. Student evaluation of teaching ratings were at least moderately valid, in that student ratings of course quality correlated positively with other measures of teaching effectiveness (Centra 1993). In one type of study, multiple sections of the same course are taught by different instructors, but there is a common final exam. The ratings instructors receive turn out to be positively correlated with the performance of their students on the exam. SETs also tend to correlate well with retrospective evaluations by alumni who rarely change their evaluations of their teachers as the years pass (Centra 1993). Other methods of evaluating teaching effectiveness do not appear to be valid. Ratings by peers and trained observers are not even reliable (a necessary condition for validity) colleagues and observers do not even substantially agree with each other in instructor ratings (Marsh and Roche 1997).

Numerous studies have cast doubt on the validity of student evaluation instrument. Rodin and Rodin (1972) found a negative relationship between student performance and student ratings. O'Connell & Dickinson (1993) found that amount learned by students was unrelated to the overall ratings of the instructor. Yunker and Yunker (2003) found that students from a class where the instructor was rated higher did worse than students from class where an instructor was rated lower, in a subsequent follow-up class. Koon and Murray (1995) found that final examination scores had only a 0.30 correlation with student ratings of instructors.
Earlier studies agree that validity of the ratings was affected by workload and students' grades (Marsh 1987). Students lacked the level of knowledge necessary to properly evaluate their instruction (Olshavsky and Spreng, 1995). Moreover the "Entertainment" level of the classroom experience has been shown to affect overall instructor ratings (Costin, Greenough, and Menges, 1971). An enthusiastic lecturer was highly rated on teaching quality despite a lecture intentionally devoid of content. Dowell and Neal (1983) concluded that the situational variables so thoroughly contaminate the validity of self□reported student learning and teacher effectiveness indices that they can only be regarded as indices of "consumer satisfaction". Cross discipline ratings bias makes the validity of SETs questionable, Cashin (1990) examined very large databases of students' ratings and found sizable differences in how students rate teaching across various academic disciplines. The variables that distinguish a required course from an elective, and that identify courses by level (freshman, second year etc) have generated significant differences in student ratings (Aleamoni 1989).
Student characteristics, instructor characteristics and class characteristics yield internally invalid and uninterpretable estimates of rating validity (Damron 1996). This challenges the accuracy and developmental utility of student ratings. Damron (1996) adds that even if a sufficiently valid rating questionnaire existed, there are no guarantees that interpretations of ratings data will be valid or consistent. Franklin and Theall (1990) noted that the users may be unskilled and may make decisions based on invalid interpretations of ambiguous or bad data. Besides ratings are particularly subject to sampling problems. Emery et al (2003) opine that errors that may render the interpretation invalid may arise due to misinterpretation of statistics which could lead to a decision favoring one instructor over another, when in fact the two instructors may not be significantly different. Even when data is adequate there may be failure to distinguish significant differences from insignificant ones. Failure to use data from available reports may be prejudicial to an instructor whose performance has been outstanding but whom, as a result of the error, is not appropriately rewarded or is penalized. Also given significant differences, there is a failure to account for or correctly identify the sources of differences for example a personal predisposition toward teaching style may lead a user to attribute negative meanings to good ratings, or to misinterpret the results of an item as negative evidence when the item is actually irrelevant and there is no quantitative justification for such a decision (Franklin and Theall, 1990). Lecture content was found to have a sizable influence on student achievement, but only a negligible impact on student ratings. Ratings, which are frequently used to make tenure and promotion decisions, were particularly elevated by instructor expressiveness. Student instructional ratings should not be used to make decisions on faculty promotion and tenure, because they are based on lecturer characteristics (e.g. charismatic and enthusiastic) rather than student outcomes (Abrami et al., 1982; Damron 1996).
According to Narcisse and Harcourt (2008) fairness perceptions are of three main types. First, distributive justice refers to the perceived fairness of an actual appraisal rating. Second, procedural justice refers to the perceived fairness of procedures used to determine the appraisal rating. Third, interactional justice refers to the perceived fairness of the rater's interpersonal treatment of the rate during the appraisal process. Overall, results show that distributive, procedural, and interactional justice factors influence employee perceptions of fairness in their appraisals.

2.2.4. Perception towards Sets

Some lecturers argue that student evaluation of teaching are a threat to academic freedom (Haskel 1997), influence instructors' grading policies, teaching style, and course difficulty, and may also restrict what a professor says in class as they may feel inhibited from discussing controversial ideas or challenging students' beliefs, for fear that some students will express their disagreement through the course evaluation form. Professors are forced to think like politicians, seeking to avoid giving offense and putting style before substance (William and Ceci 1997).
Urevbu (1997) opines that many lecturers consider student evaluation as abnormal since it means giving students a voice yet they are not competent or mature enough to evaluate lecturers. Cross (2002) suggests that SET does little general good and some particular harm to individual academics though it may be good for university education. It may arouse unhealthy competition among faculty members according to Richmond (2003). According to Iyamu (1998) rating on a five point scale at the end of the semester cannot effectively measure accurately the complexity and multidimensionality of teaching. Scholars have not arrived at a unanimous agreement on what is effective teaching. Ede and Sam (2005) students rating may be misleading, inaccurate as students are not well trained to evaluate lecturers. Students have limited background and knowledge and should not evaluate instructors' knowledge of subject matter, learning objectives, grading standards which should be judged by peers.
Emery (2001) argued that SETs are inappropriate and students should not evaluate lecturers as they should be viewed as the products of the program rather than its customers. Lecturers are the immediate customers and the industry/society the ultimate customer. From this position, it is clear that the use of SETs, which implicitly captures lecturer popularity, is inappropriate for measuring instructional effectiveness (i.e. learning). Ironically, while business departments purport to use student appraisals to increase total quality, Deming (1986) has suggested that the practice is inaccurate and demoralizing.

2.3. Research Gaps

Though student evaluation of lecturers is commonly practiced in Kenyan universities, the researcher has not come across research exploring perception and the factors that influence that perception towards student based evaluation of teaching.

3. Research Methodology

3.1. Research Design

The research design adopted by the researcher was survey or descriptive. This design was adopted for this study because it involves describing lecturer perception of student evaluation of teaching without influencing it in any way (Bell 2010). According to Shaughnessy (2011) surveys consists of a predetermined set of questions given to a sample. With a representative sample, one can describe the attitudes of the population from which the sample was drawn. Further, one can compare the attitudes of different populations as well as look for changes in attitudes over time. It also allows one to generalize the findings from the sample to the population.

3.2. Target Population

The target population for a survey is the entire set of units for which the survey data are to be used to make inferences. Thus, the target population defines those units for which the findings of the survey are meant to generalize. In this research the target population was 1114 full time academic staff from two public and two private universities within the Republic of Kenya (CHE 2011). The universities included Kenyatta University, Masinde Muliro University of science and Technology, Daystar University and Mount Kenya University. In each category one university is relatively old while the other is relatively young to establish whether perceptions might vary.

3.3. Sample Size

Sample size was obtained using the following formula, Taro (1973) where

N= Population size, n = sample size, P is the degree of variability (0.5) and ∈ is the sampling error or level of precision expressed in percentage (5 % or 0.05).

$$n = \frac{N}{1 + N(\in)^2}$$

$$n = \frac{1114}{1 + 1114\,(0.05)^2} = 294$$

Sample size was distributed as follows, Kenyatta University (185), Masinde Muliro (62), Mt Kenya (26) and Daystar (21).

3.4. Sampling Method

Universities were stratified into private and public to constitute two sub-groups after which each stratum was sampled as an independent sub-population out of which individual elements were selected randomly (Groves et al. 2009). The researcher selected this method because the sub-groups were homogenous. The strata should be mutually exclusive and every element in the population must be assigned to only one stratum. Subgroups were proportional to the population size obtained by selecting subjects so that sub-groups percentages in the population were reflected in the sample (Kombo & Tromp 2006).The universities in each stratum were selected using stratified purposive sampling, a non-probability approach based on age to ensure one relatively 'old' and relatively 'young' is selected to establish whether being new would differ or be parallel with the established ones in terms of competitiveness. Schools in each university were also similarly selected to ensure representation from social sciences, sciences, education and business. This approach as recommended by Paton (1990) illustrates characteristics of particular sub-groups of interest and facilitates comparison between different groups. Simple random sampling was finally used to select full time lecturers as respondents from each school. A complete list of all the lecturers was made and a number assigned to each of them. A set of random numbers, which identified the sample size to be sampled, was drawn. This approach gave every lecturer in the department an equal chance of being selected and gives the same characteristics and composition as the population (Kothari 2003). Sampling was without replacement and each element was sampled only once.

3.5. Data Collection

Primary and secondary sources of data were used in this research. A questionnaire was designed and administered to the academic staff. It was chosen as it provides a more comprehensive view than any other research tool and is able to collect data from a large number of respondents (Kombo & Tromp 2006). It allows the researcher to control and focus responses to the research objectives thus, enhancing relevance of data collected. They are also easy to analyze and most statistical analysis software such as SPSS can be used to process them.

3.6. Data Collection Procedure

The researcher with the help of assistants delivered the questionnaires to the sampled schools and issued to the respondents. The questionnaires were collected on the same day or on appointment within the period of data collection through the office of the dean. This procedure is economical in time and resources.

4. Data Processing and Analysis

4.1. Response rate

There were sample size was 294 respondents from the public and private universities. A total of 182 duly filled questionnaires were returned and used for this study, 143 from private and 39 from private universities making up a 61.9% response rate close enough to the recommended response rate of approximately 60% which should be the goal of researchers (Draugalis et al. 2008).

	Frequency	Percent	Valid Percent
Public university	143	78.6	78.6
Private university	39	21.4	21.4
Total	182	100.0	100.0

Table 1: Response by University

4.2. Demographics

4.2.1. Distribution by Gender

Distribution by gender was also factored into consideration and there were more males at 96 (52.7%) and females being a total of 86 (47.3%). This presents a balanced representation between the two genders. This is in line with the study to provide quantitative data from a cross section of the chosen population.

		Frequency	Percent	Valid Percent
Valid	Male	96	52.7	52.7
	Female	86	47.3	47.3
	Total	182	100.0	100.0

Table 2: Response by Gender

4.2.2. Distribution by Academic Status

Majority of the respondents were holders of a masters degree (65.4%), PhD holders accounted for 20.3 %, while those pursuing PhD were 12.1% and only 2% were holders of bachelor's degree.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	PhD	37	20.3	20.3	20.3
	Bachelors	4	2.2	2.2	22.5
	Masters	119	65.4	65.4	87.9
	PhD Candidate	22	12.1	12.1	100.0
	Total	182	100.0	100.0	

Table 3: Academic status

4.3. Measures of Student Evaluation of Teaching

Respondents were asked to rate the importance of various measures of student evaluation. The results indicate that course content, student grading and content delivery were highly considered most important by students when grading their lecturers. Time management, lecturer confidence and flexibility were not considered important by students. The findings are shown in table 4 showing frequencies and percentages.

Variable	Most important	Important	Neither important nor important	Unimportant	Highly unimportant
Course content	140(76.9%)	39(21.4%)	3 (1.6%)	0	0
Student Grading	100(54.9%)	79(43.4%)	2 (1.1%)	1 (0.5%)	0
Content delivery	63(34.6%)	97(53.3%)	21 (11.5%)	1 (0.5%)	0
Time management	12(6.6%)	96(52.7%)	62 (34.1%)	12 (6.6%)	0
Lecturer confidence	8 (4.4%)	69(37.9%)	58 (31.9 %)	30 (16.5%)	17 (9.3%)

Lecturer Flexibility	0	39(21.4%)	58 (31.9%)	40 (22.0%)	45 (24.7%)

Table 4: Frequencies and percentages of importance students attached to measures of SETs

4.4. Relationships between Independent and Dependent Variables

4.4.1. Influence of Gender on Lecturer perception to student evaluation of teaching

No	Item	Gender	SA	A	N	D	SD
1	SETs is gender biased against female academics	Male	1(1.0%)	14(14.6%)	46(47.9%)	24(25.0%)	11(11.5%)
		Female	27(31.4%)	56(65.1)	3(3.5%)	0	0
2	Students take female lecturers less seriously than males	Male	2(2.1%)	14(14.6%)	51(53.1%)	24(25.0%)	5(5.2%)
		Female	17(19.8%)	56(65.1%)	12(14.0%)	0	1.2(4.0%)
3	Men are intellectually credible than females	Male	4(4.2)	22(22.9%)	48(50.0%)	13(13.5%)	9(9.4%)
		Female	1(1.2%)	14(16.3%)	7(8.1%)	32(37.2%)	32(37.2%)
4	Women academics must work harder to prove themselves academically than males	Male	0	14(14.6%)	61(63.5%)	20(20.8%)	1(1.0%)
		Female	10(11.6%)	61(70.9%)	13(70.9%)	2(2.3%)	0
5	Males students are likely to challenge lecturers authority than female students	Male	0	16(16.7%)	64(66.7%	16(16.7%)	0
		Female	6(7.0%)	68(79.1%)	12(14.0%)	0	0
6	Female academics are expected to play nurturing role to students	Male	2(2.1%)	22(22.9%)	58(60.4%)	14(14.6%)	
		Female	4(4.7%)	66(76.7%)	16(18.6%)	0	0
7	Students' judgment about women academics capability seems to be contaminated by judgment about femininity	Male	1(1.0%)	22(22.9%)	55(57.3%)	17(17.7%)	1(1.0%)
		Female	7(8.1%)	63(73.3%)	16(18.6%)	0	0
8	Students constantly criticize female academics	Male	1(1.0%)	17(17.7%)	63(65.6%)	14(14.6%)	1(1.0)
		Female	10(11.6%)	58(67.4%)	18(20.9%)	0	0

Table 5: Perception of gender based constructs-Frequency and Percentage

Cross tabulation was used to explore how gender was related to the various gender related constructs shown in table 5. The findings indicate that a greater percentage of female lecturers agreed with the gender related statements as follows; SETs is gender biased against female academics and students take female lecturers less seriously than males (65%), women academics must work harder to prove themselves academically than males (70.9%), males students are likely to challenge lecturers authority than female students (79%), female academics are expected to play nurturing role to students (76.7%), students' judgment about women academics capability seems to be contaminated by judgment about femininity (73%) and students constantly criticize female academics (67.4%). This shows that female academics were more sensitive to gender related constructs. However, 74% of the female academics disagreed with the notion that men are intellectually credible than females.

On all the gender related constructs, majority of male academics indicated a neutral stand (neither agree nor disagree); Males students are likely to challenge lecturers authority than female students (66.6%), Students constantly criticize female academics (65.6%), Female academics are expected to play nurturing role to students (60.4%), Women academics must work harder to prove themselves academically than males (63.5%), Students take female lecturers less seriously than males (53.1%). This indicates that male academics were less sensitive to gender based issues than females. Gender therefore is a strong determinant of perception towards SETs.

Earlier studies by Kirkpatrick (1997) opined that female lecturers were more sensitive to the harm that SETs would inflict on them than their male counterparts. Gender was found to have significant effect on lecturer perception (Harun, et al. 2011), an opinion supported by Abdulrahim (2010). However Idaka et al. (2006) and Olatoye (2011) disagreed with this view and found gender to have less significant effect. Teachers in most American schools were disposed towards student evaluation both males and females (Smith and Anderson 2003). Gender variances were found in the cognitive bases of employee work□oriented attitudes and these were reflected through measures of perceptions of the utility and relevance of formal organizational appraisal systems. Overall, the results indicated that females and males use different information bases when evaluating performance appraisal systems (Hind and Baruch 1997). As the probability increases that employee performance is evaluated by a female, women expect more positive

outcomes of subjective, but not objective evaluation processes done by fellow females (Maas and Gonzalez 2011). According to Kashane (2009) aspects such as ethics, fairness, motivation, accuracy, validity, rating errors, effectiveness and feedback, should therefore be examined in more detail in order to determine where specific problem areas may lie between males and females. The results indicated statistically significant differences in perception between males and females in terms of fairness, motivation, and feedback.

4.4.2. Effect of Uses of Student Evaluation of Teaching on Lecturer Perception

Lecturers were asked to indicate their level of agreement with the various uses that SETs are put into. The uses are classified into summative items (1-5) and formative (items 6-12). Most academic staff were dissatisfied with the use of SETs for summative purposes with mean ranging from 1.6374 and 94% (SETs should be used to make pay decisions) to 2.2692 and 67% (SETs should be used for administrative decisions). With regard to formative uses, majority of respondents agreed that SETs improve lecturers teaching, 73.1% (mean 3.93), improve classroom instruction ,72% (mean 3.91) and enhancing lecturers' professional growth and improving student-lecturer relationship at 58.2% (3.7088) and 53.8% (3.7473) respectively as summarized in table 6.

No	Item	SA	A	N	D	SD	Mean
1	SETs should be used to make pay decisions	5 (2.7)	3 (1.6)	1 (0.5)	85 (46.7)	88 (48.4)	1.6374
2	SETs should be used to determine promotion	5(2.7)	3(1.6)	2(1.1)	100(54.9)	72(39.6)	1.7308
3	SETs should be used to determine lecturer retention	4(2.2)	7(3.8)	2(1.1)	124(68.1)	45(24.7)	1.9066
4	SETs should be used to determine faculty awards	0	15(8.2)	44(24.2)	102(56)	21(11.5)	2.2912
5	SETs should be used for administrative decisions	0	13(7.1)	46(25.3)	100(54.9)	23(12.6)	2.2692
6	SETs is used to improve lecturers teaching	19(10.4)	133(73.1)	26(14.3)	4(2.2)	0	3.9176
7	SETs improve classroom instruction	20(11%)	131(72)	30(16.5)	1(0.5)	0	3.9341
8	Enhances lecturers' professional growth	17(9.3)	106(58.2)	48(26.4)	11(6)	0	3.7088
9	Enhances Self Evaluation	4(2.2)	92(50.5)	77(42.3)	9(4.9)	0	3.6000
10	Improves Student-Lecturer relationship	21(11.5)	98(53.8)	59(32.4)	4(2.2)	0	3.7473
11	Helps lecturers become more innovative	2(1.1)	85(46.7)	81(44.5)	13(7.1)	1(0.5)	3.4066
12	Lecturers become more committed	3(1.6)	96(52.7)	74(40.7)	8(4.1)	1(0.5)	3.5055

Table 6: Uses/Purposes of Student Evaluation of teaching (Frequency, Percentage and Mean)

Penny (2003) opines that the evaluation reports are used for promotional exercises and strategic decisions such as retention, tenure, curriculum development, external quality care and research on teaching. Under formative uses, SETS is used for classroom improvement of classroom instruction, student learning and fosters growth of the lecturer. Malgul (1998) felt that evaluation when properly executed disclosed what skills and knowledge lectures have brought vis a vis the skills and knowledge needed to meet the demands of producing effective students. It helps identifies deficiencies and strengths in HR performance and helps staff improve their teaching. According to Hoyton and Pallet (1999) SET scores were used to make comparisons between lecturers in different departments and faculties. Isiaka (1998) adds that lecturers were positively disposed towards SETs for formative purposes only a view supported by Machingambi and Wadesango (2011) who found that lecturers in South African universities were positively disposed towards student evaluation for formative purposes. According to Samian and Noor (2012) the main purpose of student evaluation of teaching is to provide information and feedback to academic staff on the teaching performance which in turn should enable them to do some self reflection and eventually take necessary actions to enhance their teaching performance in the future. Kulik (2001) in his research suggests that SETs as used to improve teaching effectiveness and monitoring performance of graduate assistants improve teaching skills and document these skills when applying for promotions.

Summative uses of evaluation according to Gold (2001) include promotion, pay decisions, demotion dismissal and awards. Lecturers in South Africa and Nigeria were found to be negatively disposed towards student evaluation for summative purposes

(Iyamu and Aduwa 2005, Machingambi and Wadesango 2011). Abrami (1982) argue that SETs should not be used for summative purpose because charismatic and enthusiastic lecturers receive favourable ratings regardless of how well they know their subject matter.

4.4.3. Effect of Situational factors (validity and reliability) on lecturers' Perception

Respondents were asked to indicate their level of agreement with various situational factors relating to validity and reliability of SETs. There was unanimous disagreement with the assertion that SETs are fair valid and reliable (items 1 and 2) with most expressing disagreement (43-46%) and neutral stand (28-31%). Mean ranged from 2.1 to 2.2.

Majority of respondents (50%) agreed that students may use SETs to get even with lecturers they dislike and rated favourably lecturers who graded them highly (54% and 34%) mean of 4.4670. A high percentage also agreed that students from large classes rated academics less favourably (45% to 49%) and that SETs contained irrelevant variables (41% and 53%). These responses cast aspersions on the validity of SETs as the summary in table 7 indicates.

No	Item	SA	A	N	D	SD	Mean
1	SETs are fair method of evaluation	0	11(6.0%)	57(31.3%)	79(43.4%)	35(19.2%)	2.2418
2	SETs are reliable and valid	0	4(2.2%)	52(28.6%)	85(46.7%)	41(22.5%)	2.1044
3	Students use SETs to get even with lecturers they dislike	14(7.7%)	91(50%)	68(37.4%)	8(4.4%)	1(0.5%)	3.5989
4	Students ratings reflect teaching effectiveness	3(1.6%)	46(25.3%)	105(57.7%)	17(9.3%)	10(5.5%)	3.0659
5	Students rate favourably lecturers who grade them highly	100(54.9)	68(37.4%)	13(7.1%)	1(.5%)	0	4.4670
6	Students possess skills and knowledge to evaluate lecturers	0	23(12.6%)	66(36.3%)	78(42.9%)	15(8.2%)	2.5330
7	Using SETs across various disciplines reduces their validity	10(5.5%)	112(61.5)	58(31.9%)	2(1.1%)	0	3.7143
8	Students from large classes rate lecturers less favourably	2(1.1%)	90(49.5%)	83(45.6%)	7(3.8%)	0	3.4780
9	SETs contain irrelevant variables	3(1.6%)	98(53.8%)	76(41.8%)	5(2.7%)	0	3.5440

Table 7: Effect of Situational Factors on Perception to SETs

Braskamp and Ory (1994) found that lecturers were skeptical about SETs as it might damage their careers and were not valid or reliable. SETs lack of validity could be caused by user error, rater qualification error and defamation, since students are not qualified to evaluate lecturers, the lack of validity causes lecturers to dismiss the evaluation tool. According to Iyamu and Aduwa (2005) due to the importance of teacher evaluation it is imperative that evaluation instruments used be both valid and reliable. Stoklasa et al. (2011) found that lecturers in Hungary did not consider student ratings as valid measure of teaching effectiveness and is more or less a popularity contest. Makondo and Ndebele (2014) found that some SETs had non-applicable questions unsuitable for some disciplines and the instrument is more on traditional teaching. Students are not qualified to evaluate their lecturers, student ratings lack a certain degree of behavioural specificity (i.e. a five□point Likert scale) (Cascio and Bernardin, 1981). No meaningful estimate of the validity of student ratings can be provided with confidence that is generalizable enough to be useful (Dowell and Neal, 1982).

Student knowledge of how the results of the evaluation (salary, promotion, tenure consideration) will be used tends to produce favourable ratings than if the purpose was for lecturer feedback and course improvement (Makondo and Ndebele 2014). Validity was also affected by the grades given to the students where those happy about their grade rewarded the lecturer with favourable rating and vice versa. They also found that lecturers felt students were biased where classes were large and new lecturers felt the evaluation was not fair to them as students did not have rapport with them. In his research of higher education institution in Iran, Shirbagi (2007) reported that how much students perceive the lecturer as being charismatic is an important predictor of SET scores.

According to Goksoy and Alayoglu(2013) ethics in decision making has been an issue for academics and practitioners, perception of performance evaluation fairness has an impact on employees' ethical decision making. Choom and Embi (2012) add that in an organization, fair performance evaluation is essential to ensure that no victim will be harmed or purposely being harmed. Generally, subjectivity tends to be a major unfair element causing the unfairness in the performance evaluation. Kulik (2001) in his research came across lecturers who felt that student ratings were reliable and valid measures that brought scientific accuracy to

the evaluation of teaching and gave students a voice in shaping the goals and practices of their university giving them a sense of participation.

4.4.4. Dependent variable- Perception of lecturers towards student based evaluation

Variable	SA	A	N	D	SD	Mean
It is acceptable for students to evaluate lecturers	1 (0.5%)	24(13.2%)	42(23.1%)	75 (41.2 %)	40(22.0%)	2.2912
SETs improve student lecturer interaction	1 (0.5%)	38(20.9%)	71(39.6%)	66 (36.3%)	6 (3.3%)	2.7912
SETs provide direct observation of lecturer teaching	0	72(39.6%)	57(31.3%)	53 (29.1%)	0	3.1044
A recognition of student involvement in shaping university goals and practice	12 (6.6%)	87(47.8%)	38 (20.9%)	45(24.7%)	0	3.3623
Increases chances of excelling in teaching	0	44(24.2%)	78 (42.9 %)	59 (32.4%)	1 (0.5%)	2.9066
Help lecturers be more accountable	0	41(22.5%)	82 (45.1%)	56 (30.8%)	3 (1.6%)	2.8846

Table 8: Perception of SETs

Respondents were asked to indicate their level of agreement with various constructs related to perception on SETs. More than half of the academics (72%) indicated their disagreement with the use of SETs to evaluate lecturers and 63% that SETs improve student lecturer interaction. 48% agreed that SETs were recognition of student involvement in shaping university goals and practice. Only 40% agreed that SETs provided direct observation of lecturer teaching and with regard to whether student based evaluation increases chances of excelling in teaching and, helps lecturers be more accountable majority of respondents (42-45%) took a neutral stand.

Student evaluation is the key performance index for lecturers in the staff appraisal and teaching effectiveness. Lectures performance contributes to student satisfaction which in turn affects university image and student loyalty (Helgessen and Nesset 2007). Lecturers in Ghana and Kenyan colleges accepted the idea of student evaluation of classroom effectiveness (Isiaka 1998). Lecturers felt that students have a right to make judgment about quality of teaching (Ede and Sam (2005. Iyamu and Aduwa (2005) and Imogie (2000) found that more experienced lecturers were found to show higher preference for students rating than junior lecturers. According to Richmond (2003) student opinion is important as it gives additional data based on direct observation that is used to judge lecturers competence. David and Adebowale (1997) found that SETs increases chances of recognizing and rewarding excellence in teaching, provides means of interaction between the teacher and student, provides direct and extensive information about the lecturer, provides tangible evidence of student recognition and involvement, and can be used to improve classroom instruction, student learning and foster professional growth of the lecturer.

David and Adeboale (1997) came across lecturers who were skeptical about SETs. Richmond (2002) found that though student evaluation is important lecturers had negative perceptions towards it. SETs were found to be reliable and stable but could be affected by potentially biased students due to their emotional states (Chonko et al. 2002) as positive emotions help students assess situations more rationally while Pekrun et al. (2002) add that negative emotions can inhibit students evaluations. Kulik (2001) found lecturers who felt that ratings used in SET were confusing and inconclusive, gave meaningless quantification and gave students a chance to use them to get even with their lecturers. Besides SETs are about personality contest measure than assessment of teaching effectiveness (Sproule 2002). According to Beran and Rokosh (2009) lecturers felt the SETs were poorly designed to be useful in improving teaching practices as their quality and legitimacy was compromised and therefore perceived them negatively.

4.5. Hypotheses Testing on Relationships between Independent and Dependent variables

The first hypothesis was tested using Independent Samples Test due to the need to compare mean scores for male and female respondents.

- H0: There is no positive relationship between gender and lecturer perception of student evaluation of teaching
- H1: There is a positive relationship between gender and lecturer perception of student evaluation of teaching

From the resulting statistics of the t-test the mean scores indicate that males scored least on the following factors, SETs is gender biased against female academics (2.6875), Students take female lecturers less seriously than males (2.8333) while females scored highest in the two factors at 4.2791 and 4.0349 respectively. In all other factors female scores were higher than those of males.

	Gender	N	Mean	Std. Deviation	Std. Error Mean
SETs is gender biased against female academics	Male	96	2.6875	.89810	.09166
	Female	86	4.2791	.52359	.05646
Students take female lecturers less seriously than males	Male Female	96	2.8333	.81650	.08333
		86	4.0349	.62210	.06708
Men are intellectually credible than females	Male	96	2.9896	.95691	.09766
	Female	86	2.0698	1.10390	.11904
Women academics must work harder to prove themselves academically than males	Male Female	96	2.9167	.62688	.06398
		86	3.9186	.59833	.06452
Males students are likely to challenge lecturers authority than female students	Male Female	96	3.0000	.58038	.05923
		86	3.9302	.45480	.04904
Female academics are expected to play nurturing role to students	Male Female	96	3.1250	.66886	.06826
		86	3.8605	.46432	.05007
Students' judgment about women academics capability seems to be contaminated by judgment about femininity	Male Female	96	3.0521	.70142	.07159
		86	3.8953	.50942	.05493
Students constantly criticize female academics	Male	96	3.0313	.63995	.06531
	Female	86	3.9070	.56627	.06106

Table 9: Results of t-test comparing gender based factor scores of males and females

Gender based factors	Levene's Test for Equality of Variances			t-test for Equality of Means		
	F	Sig.	t	df	Sig.2tailed	Significant.
SETs is gender biased against female academics	20.273	.000	-14.387	180	.000	Yes
			-14.784	155.711	.000	
Students take female lecturers less seriously than males	8.647	.004	-11.068	180	.000	Yes
			-11.232	175.601	.000	
Men are intellectually credible than females	3.918	.049	6.021	180	.000	Yes
			5.974	169.308	.000	
Women academics must work harder to prove themselves academically than males	.630	.428	-10.998	180	.000	No
			-11.027	179.265	.000	

	F	Sig.	t	df	Sig.2tailed	Significant.
Males students are likely to challenge lecturers authority than female students	1.340	.249	-11.938	180	.000	No
			-12.096	176.951	.000	
Female academics are expected to play nurturing role to students	6.888	.009	-8.521	180	.000	Yes
			-8.687	169.788	.000	
Students' judgment about women academics capability seems to be contaminated by judgment about femininity	4.373	.038	-9.187	180	.000	Yes
			-9.345	172.837	.000	
Students constantly criticize female academics	.004	.949	-9.729	180	.000	No
			-9.794	179.975	.000	

Table 10: Results of T-test

There was significant difference in scores for males and females in the following constructs: SETs is gender biased against female academics (t= -14.387, p<0.05), students take female lecturers less seriously than males (t= -11.068, p<0.05), men are intellectually credible than females (t= 6.021, p<0.05), female academics are expected to play nurturing role to students (t= -8.521, p<0.05), and students' judgment about women academics capability seems to be contaminated by judgment about femininity (t= -9.187, p<0.05). There was no significant difference in scores for: males and females in women academics must work harder to prove themselves academically than males (t= -10.998, p> 0.005), Males students are likely to challenge lecturers authority than female students (t= -11.938, p>0.05), Students constantly criticize female academics (t= -9.729, p> 0.05). Since the significance level is less than the p-value in majority of the constructs it can be concluded that there is a significant relationship between gender and perception of lecturers towards student based evaluation therefore the null hypothesis is rejected.

The next hypotheses were tested using f-test

- H0: There is no positive relationship between summative uses of SETs and lecturer perception of student evaluation of teaching
- H0: There is no positive relationship between formative uses of SETs and lecturer perception of student evaluation of teaching
- H1: There is a positive relationship between summative uses of SETs and lecturer perception of student evaluation of teaching
- H1: There is a positive relationship between formative uses of SETs and lecturer perception of student evaluation of teaching

The proportion of variance in the dependent variable is 0.22% (R square). Use of SETs for summative purposes does not statistically significantly predict perception to SETs since F (1,182) = 4.127, p=044 > 0.01. Since the significance level 0.044 is greater than (>) the p-value 0.01 the null hypothesis is not rejected.

The proportion of variance in the dependent variable is 0.35% (R square). Use of SETs for formative purposes does not have a statistically significant relationship to perception of academic staff to SETs since F (1,182) = 6.540, p=011 < 0.05. Since the significance level 0.011 is less than (<) the p-value 0.05 the null hypothesis is rejected. In the case of formative and summative uses of SETs, they have a positive relationship with perception to SETs.

- H0: There is no positive relationship between situational factors and lecturer perception of student evaluation of teaching.
- H1: There is a positive relationship between situational factors and lecturer perception of student evaluation of teaching.

The proportion of variance in the dependent variable is 0 % (R square). Situational factors (validity and reliability) does not have a statistically significant relationship to perception of academic staff to SETs since F (1,182) = .037, p=.848 >0.05. Since the significance level 0.848 is greater than the p-value 0.05 the null hypothesis is not rejected therefore situational factors do not have a strong effect of perception of lecturers towards student based evaluation.

4.6. Summary

Female academics were more sensitive to gender related construct while male academics were less sensitive to gender based issues. Gender therefore is a strong determinant of perception towards SETs. There is a positive relationship between gender and lecturer perception of student evaluation of teaching. Most academic staff were dissatisfied with the use of SETs for summative

purposes with mean ranging from 1.6374 to 2.2692 while on the other hand, use of SETs for formative purposes was agreeable to most academic staff with means ranging from 3.4 to 3.9.There was unanimous disagreement with the assertion that SETs are fair valid and reliable with most expressing disagreement (43-46%) and neutral stand (28-31%). Mean ranged from 2.1 to 2.2.. More than half of the academics (72%) indicated their disagreement with the use of SETs to evaluate lecturers indicating a negative perception by academic staff towards student based evaluation of teaching.

4.7. Conclusion

There is a positive relationship between gender and lecturer perception of student evaluation of teaching. Most academic staff was dissatisfied with the use of SETs for summative purposes while on the other hand, use of SETs for formative purposes was agreeable to most academic staff. Most academic staff felt that student based evaluation of teaching were not fair, valid or reliable. Majority of the academics disagreed with the use of SETs for their evaluation.

4.8. Recommendations

Organizations should design and use performance evaluation methods that ensure equal opportunities for men and women since as seen in this research findings gender discrimination is still a problem in organizations and societies. It is wise for organizations to blend SETs along with other evaluation methods to increase their validity and reliability in order to gain more acceptances by the academic staff. Evaluations could be done mid semester to enable academics make changes to improve course delivery in units they were instructing and for which they were being evaluated. Ordinarily student evaluations are done towards the end of the semester and lecturers do not have the ability to make changes for the course in which they are being evaluated. Mid-semester evaluation would be more helpful for the lecturer to improve on the ongoing course unit.

5. References

1. Abdulraheem, Y. (2010). University teachers perception of the effects of student evaluation of teaching on lecturers instructional practices in Nigeria. Proceedings of the 1st international conference on collaborational of educational faculties in West Africa, Ilorin, Nigeria.
2. Abrami, P.C., Leventhal, L. & Perry, R.P. (1982). Educational seduction. Review of Educational research, 32, 446-464.
3. Aleamoni, L. (1987). Student rating: myths versus research facts. Journal of Personnel Evaluation in Education, 1, 111□9.
4. Aleamoni, L. (1989). Typical faculty concerns about evaluation of teaching. in Aleamoni, L.M. (Ed.), Techniques for Evaluating and Improving Instruction, San Francisco:Jossey□Bass.
5. Ambrose, M. L., & Kulik, C. T. (1999). Old friends, new faces: Motivation research in the 1990s. Journal of Management, 25, 231–292.
6. Ambrose, M. L., & Kulik, C. T. (1999). Old friends, new faces: Motivation research in the 1990s. Journal of Management, 25, 231–292.
7. Bagihole, B. & Woodward, H. (1995). An occupational hazard warning: academic life can seriously damage your health. An investigation of sexual harassment of women academics in a UK university. British Journal of Sociology of Education, 16, 37-51.
8. Barnes-Powell, T., & Letherby, G. (1998). All in a days' work: Gendered care work in higher education. In Malina, D. & Maslin, P.S. (eds), Surviving the Academy: Feminist perspective, London: Palmer.
9. Barrette, C. Morton, E., & Tozcu, A. (2006). An overview of teacher evaluation. Educational evaluation research and reviews, 2, 80-84.
10. Belanger, C. H. & Longden, B. (2009). The effective teachers characteristics as perceived by students. Tertiary education and Management, 15(4), 323-340.
11. Benz, C.R. & Blatt, S.J. (1996). Meanings and underlying students ratings of faculty. Review of Higher Education, 19, 411-433.
12. Beran, T.N. & Rokosh, J.L. (2009). The consequential validity of student ratings: what do instructors really think? Alberta journal of Educational Research, 55(4), 497-511.
13. Boyd, N. M., & Kyle, K. (2004). Expanding the view of performance appraisal by introducing social justice concerns, Administrative Theory & Praxis, 26 (3), 249–277.
14. Braskamp, L.A. and Ory, J.C. (1994), Assessing Faculty Work: Enhancing Individual and Institutional Performance, San Francisco: Jossey□Bass.
15. Carson, L. (2001). Gender relations in higher education: exploring lecturers perception of student evaluation of teaching. Research Papers in Education, 16(4), 337-358.
16. Cascio, W.F. & Bernardin, H.J. (1981). Implications of performance appraisal litigation for personnel decisions, Personnel Psychology, 34, 211□26.
17. Cashin, W. (1990). Students do rate different academic fields differently. In Theall, M. and Franklin J. (Eds), Student Ratings Of Instruction: Issues For Improving Practice, San Francisco:Jossey□Bass.
18. Centra J A.(2003). Will teachers receive higher standards ratings by giving higher grades and less course work. Research in higher education, 44(5)495-518.
19. Centra, J. A. (1993). Reflective Faculty Evaluation .San Francisco: Jossey-Bass Publishers.
20. Chonko, L B., Tanner,J.F., Davis, R. (2002). What are they thinking? Students expectation and self assessment. Journal of Education for business, May-June 271-281.

21. Choom L.K. & Embi, M.A. (2012).Subjectivity, Organizational Justice and Performance Appraisal: Understanding the Concept of Subjectivity in Leading Towards Employees' Perception of Fairness in the Performance Appraisal. Journal of social behavioural sciences, 62.
22. Christopher, K.S. & Shane, P.D. (2007). Pharmacy students perception of teaching evaluation process. American journal of pharmaceautical education, 71(1).
23. Costin, F., Greenough, W.T., & Menges, R.J., (1971). Student ratings of college teaching: Reliability, validity, and usefulness. Review of Educational Research. 41, 511-535.
24. Creswell, J. W. (2003). Research Design: Qualitative, Qualitative and Mixed Methods Approaches. London: Sage Publication.
25. Cross, R. (2002). Measuring Quality in Education. New York: El-Kley.
26. David and Adebowale (1997). Student evaluation of teaching effectiveness, A Nigerian Investigation. Higher Education, 24(4), 453-463.
27. Deming, W.E. (1986). Out of the Crisis, MIT Center for Advanced Engineering Study, Cambridge, MA.
28. DeNisi, A.S., & Gonzales, J.A. (2000). Design performance appraisal systems to improve performance, in Locke, E.A. (Eds), Handbook of Principles of Organizational Behavior. Oxford: Blackwell, 60-72.
29. Dowell, D.A. & Neal, J.A. (1982). A selective review of the validity of student ratings of teaching. Journal of Higher Education, 53, 51□62.
30. Dowell, D.A. & Neal, J.A. (1983). The validity and accuracy of student ratings of instruction: a reply to Peter A. Cohen. Journal of Higher Education, 54, 459□63.
31. Draugalis, JR, Coons, & S.J, Plaza, C.M. (2008). Best Practices for Survey Research Reports: A Synopsis for Authors and Reviewers. Am J Pharm Educ.(1)72. Article 11.
32. Draugalis, JR, Coons, & S.J, Plaza, C.M. (2008). Best Practices for Survey Research Reports: A Synopsis for Authors and Reviewers. Am J Pharm Educ.(1)72. Article 11.
33. Ede, O.S. & Oglebaen, S.E.A. (2005). Lecturers perception of student evaluation in Nigerian universities. International Educational journal, 6(5), 619-625.
34. Ehie, I. C. & Karathanos, D. (1994). Business faculty performance evaluation based on the new AACSB accreditation standards. Journal of Education for Business 69 (5), 257-262.
35. Emery, C.R., Kramer, T.R., & Tian, R.G. (2003). return to Academic standards: a critique of student evaluations of teaching effectiveness. Retrieved April 30th 2014 from www.emerald insight.com/research. Quality Assurance in Education, 11(1), 37-46.
36. Emery. C. R. (2001). 'Professors as customers', In Lamb et al. (eds) Great ideas for teaching marketing, 6th ed, New York: South Western college Publishing.
37. Franklin, J. & Theall, M. (1990). Communicating student ratings to decision makers: design for good practice. in Theall, M. and Franklin J. (Eds), Student Ratings of Instruction: Issues For Improving Practice, San Francisco: Jossey□Bass.
38. George, D., & Mallery, P. (2003). SPSS for Windows step by step: A simple guide and reference. 11.0 update (4th ed.). Boston: Allyn & Bacon.
39. Goksoy, A. & Alayoglu, N. (2013). The Impact of Perception of Performance Appraisal and Distributive Justice Fairness on Employees' Ethical Decision Making in Paternalist Organizational Culture. Performance. Improvement Quarterly, 26, 57–79.
40. Gold, R. (2001). Evaluation of Instruction. Educational Studies, 15(1), 31-42.
41. Gravestock, P., & Greenleaf, G.E. 2008). Student course evaluation: Research, models and Trends. Higher Education Quality Council of Ontario: Toronto.
42. Harun, S., Dazz, S.K., Saaludin, N., & Ahmad, W.S. (2011). Lecturers Perception of Student Evaluation at University of Kuala Lumpur. Proceedings on enhancing learning; Teaching and Learning conference 2011.
43. Haskell, R. E. (1997). Academic Freedom, Tenure, and Student Evaluation of Faculty: Galloping Polls in the 21st Century," Education Policy Analysis Archives 5.
44. Heckert, T., White, K., Gulinson, A., Schneider, S. (1999). Differential treatment in Academia: Do female faculty report being treated differently by students than female faculty? Psychological Reports, 85, 263-268.
45. Helgessen O., & Nesset, E. (2007). Images, satisfaction and antecedents: Drivers of student loyalty? A case study of Norwegian University College. Corporate Reputation Review, 10(1), 38-59.
46. Hind, P., & Baruch, Y.(1997). Gender variations in perceptions of performance appraisal. Women in Management Review, 12 (7), 276 – 289.
47. Hoyt, D.P. & Pallet, W.H. (1999). Appraising teaching effectiveness: Beyond student ratings (IDEA Paper No. 36). Kansas State University Centre for faculty evaluation and development.
48. Idaka, I., Joshua, M.T. & Kritsonis, W.A. (2006). Attitude of academic staff in Nigeria tertiary educational institutions to student evaluation of instruction (SEI). National forum of Educational Administration and Supervision Journal, 23(4).
49. Imogie, A.I. (2000). Do you know who is teaching your child? Benin City: University of Benin Press.
50. Isiaka BT. (1998). Teachers perception of student evaluation of teaching effectiveness. Lagos Journal of Education, 6-11.
51. Iyamu, E.O.S & Aduwa-Oglebaen, S.E. (2005). Lecturer Perception of Student Evaluation in Nigerian universities. International Educational Journal, 6(5), 619-625.
52. Kashane, K.E. (2009). Gender perceptions and experiences of performance appraisal in selected public service institutions

53. Kelly, M. (2012). Student evaluation of teaching effectiveness: considerations for Ontario universities. Discussion Paper, COU No. 866.
54. Kirkpatrick, D.L. (1977). Determining Training Needs: Four Simple and Effective Approaches. Training and Development Journal, 31 (2) 22-5.
55. Kombo, K.D., & Tromp, D.L.A. (2006). Proposal and Thesis Writing: an Introduction. Nairobi: Paulines Publication.
56. Koon, J. & Murray, H.G., (1983). Low-inference classroom teaching behaviors and student ratings of college teaching effectiveness. Journal of Educational Psychology, 75, 138-149.
57. Kothari, C. R. (2004). Research Methodology, Methods and Techniques. (2nd ed). New Delhi: New Age International.
58. Kulik JA (2001). Student ratings: validity, utility and controversy, new directions for institutional research, 27(5)9-12.
59. Maas, V. S., & Gonzalez, R. T. (2011). Subjective Performance Evaluation and Gender Discrimination. Journal of Business Ethics, 4,667-681.
60. Machingambi, S., & Wadesango, N. (2011). University lecturers' perception of student evaluation of their instructional practices. Anthropologist, 13(3), 167-174.
61. Magner, D. K., (1997) Report says standards used to evaluated research should also be used for teaching and service. The Chronicle of Higher Education, 44 (2), 18-19.
62. Makondo, L. & Ndebele, C. (2014). University lecturers views on student-lecturer evaluation. Anthropologist, 17(2), 377-386.
63. Malgul, H. W. (1998). The role of research and evaluation in survival of education in Nigeria. Journal of Quality in Education, 3, 177-184.
64. Marsh, H. (1987). Students' evaluations of university teaching: Research findings, methodological issues, and directions for future research. International Journal of Educational Research, 11 (3), 253-388.
65. Marsh, H. (1991). Multidimensional students' evaluations of teaching effectiveness: A test of alternative higher-order structures. Journal of Educational Psychology, 83 (2), 285-296.
66. Marsh, H. W. (1995). Still weighting for the right criteria to validate student evaluations of teaching in the IDEA system. Journal of Educational Psychology, 87(4), 666-679.
67. Marsh, H.W., & Roche.L. A. (1997). Making Students' Evaluations of Teaching Effectiveness Effective. American Psychologist 52, 1187-97.
68. Marsh, W. & Bailey, M. (1993). Multidimensional students evaluation of teaching effectiveness. Journal of Higher Education, 6, 1-18.
69. Miller, J. & Chamberlain, M. (2000). Women are teachers, men are professors: a study of students perception. Teaching Sociology, 28, 283-298.
70. Milmore, M., Lewis, P., Saunders, M., Thornhill, A. & Morrow, T. (2007). Strategic Human Resource Management: Contemporary Issues. Harlow: Prentice Hall.
71. Narcisse, S., & Harcourt, M. (2008). Employee fairness perceptions of performance appraisal: a Saint Lucian case study. The International Journal of Human Resource Management, 19(6), 1152-1169.
72. Ngware M.W., & Ndirangu M. (2007). An Improvement In Instructional Quality: Can Evaluation Of Teaching Effectiveness Make A Difference? Quality Assurance in Education, 13 (3), 183-201.
73. Ngware, M. W., Onsomu, E. N., & Manda, D. K.(2005). Private Sector Investment in Education and Training: A Case of Tertiary Education in Kenya. The Kenya Institute for Public Policy Research and Analysis (KIPPRA), Special Report 7.
74. Nunnaly, J.C. (1978). Psychometric Theory. New York: McGraw Hill.
75. O'Connell, D. Q. & Dickinson, D. J., (1993). Student ratings of instruction as a function of testing conditions and perceptions of amount learned. Journal of Research and Development in Education 27, (1), 18-23.
76. Olatoye, R.A. Aanu, E.M. (2011). Senior secondary school science teachers perception of using students to evaluate teaching effectiveness. Journal of emerging trends in education research and policy studies, 164-170.
77. Olshavsky, R. W. & Spreng, R.A. (1995). Consumer satisfaction and students: Some pitfalls of being customer driven. Journal of Consumer Satisfaction Dissatisfaction and Complaining Behaviors, 8, 69-77.
78. Pekrun, R., Goetz, T., Titz,W., & Perry, R.P. (2002). Academic emotions in students self regulated learning and achievement. A program of qualitative and quantitative research. Educational psychologist, 2(37), 91-105.
79. Penny AR. (2003). Changing the agenda for research students view about teaching: four shortcomings of SRT research . Teaching in higher education, 8(3), 399-411.
80. Richmond, E. (2003). Looking at good teaching. Educational Evaluation, 35(1), 48-59.
81. Rodin, M. & Rodin, B. (1972). Student evaluations of teachers Science. 177, 1164-1166.
82. Samian, Y. & Noor,N.M. (2012). Students' perception of good lecturer based on lecturer performance assessment. Centre for teaching and learning at university technology of Malaysia. Social and behavioural science, 56,786-790.
83. Santiago P & Benavides F. (2009). Teacher evaluation, a conceptual framework and examples of country practices. Paper presented at the OECD-Mexico workshop towards a teacher evaluation framework, Mexico, 12 August, 2009.
84. Seldin P. (1993). The use and abuse of student rating of professors. The Chronicle of Higher Education, 40(1), 40.
85. Shahzad, K., Bashir, S., & Ramay, M. (2008). Impact of human resource practices on perceived performance of university teachers in Pakistan. International Review of Business Research Papers, 4 (2)302-305.
86. Shaughnessy, J., Zechmeister, E.& Jeanne, Z. (2011). Research methods in psychology (9th ed.). New York, NY: McGraw Hill.

87. Shirbagi, N. (2007). Are students evaluation of teaching valid? Evidence from an Iranian higher education institution. Bulletin of Education and Research, 29(2), 21-32.
88. Spooren, P. & Mortelmans, D. (2006). Teacher professionalism and student evaluation of teaching: will better teachers receive higher ratings and will better students give higher ratings? Educational Studies, 32(20, 201-214.
89. Sproule, R. (2002). The underdetermination of instructor performance by data from the student evaluation of teaching. Economic of Educational review, 21, 287-294.
90. Stoklasa, J., Talasova, J., & Holecek, P. (2011). Academic staff performance evaluation- variants of models. Acta Polytechnica Hungarica, 8(3).
91. Tagamori, H.T., & Bishop L.A. (1995). Student evaluation of teaching: flows in the instruments. Thoughts and Action, 11(1), 63-78.
92. Taro, Y. (1973). Statistics, an Introductory Analysis, (3rd ed). New York: Harper and Row.
93. Urevbu, A. O. (1997). Creating the schools we deserve: reflections on education, pedagogy and curriculum. Benin City: University of Benin.
94. Weinberg, BA., Fleisher, MB., & Hashimoto, M. (2007). Evaluating methods for evaluating Instructors. The case of Higher Education National Bureau of Economic Research, Working Paper: Cambridge.
95. Williams, W. M. & Ceci. S.J. (1997). How'm I Doing? Problems with Student Ratings of Instructors and Courses Change. The Magazine of Higher Learning 29, 12-23.
96. Yoder, J.D. & Schleicher, T.L. (1996). Undergraduate regard deviation from occupational gender stereotypes as costly for women. Sex Roles, 34,171-188.
97. Yosof, M.Z.M., Hamid NHA, Rahman, MA & Jaafar NH (2012). Can students emotional states influence their assessment of lecturers' performance? International Journal of Teaching and Learning, 4(1), 96-105.
98. Yunker, P. & Yunker, J. (2003). Are student evaluations of teaching valid? Evidence from an analytical business core course. Journal of Education for Business. 78, 313-317