www.jsser.org

Journal of Social Studies Education Research

Sosyal Bilgiler Eğitimi Araştırmaları Dergisi

2017:8 (3),149-167

Major Factors Affecting Students' Perception Towards Faculty Evaluation of Teaching (SET)

Smitha Dev[1] *& Nudrat Qayyum*[2]*,*

Abstract

Gathering students' feedback and evaluation during the end of every semester is believed to be valuable for instructors' growth and development. The present study is aimed to find out the validity of students' responses and various factors involved in their rating towards faculty. A total of 150 students from 1st to 4th year of Abu Dhabi University participated in this study. Both male and female students from all the disciplines contributed in the study. The study also focused on covariates like gender, age, specializations, and students' GPAs to have any relation with their rating towards professors. Pearson's correlation analysis and multiple regression analysis were employed to understand the various factors impacting students' perception of their instructors. The results revealed that gender, age, students' GPAs, and instructors' nationalities had positive effects on students' evaluation of their instructors.

Keywords: *Student evaluation, Perception, teaching practices, multidimensional, grade inflation*

Introduction

The procedure of students' evaluation of teaching (SET) at the end of semester is a norm in all of the American universities and others that share American curricula and teaching practices. In fact, it is the most frequently used method of evaluating teaching the world over (Newton, 1988; Seldin, 1989; Stratton, 1990; Badri et al., 2006). This process, with its merits and demerits, has long been a matter of criticism by many scholars and educationists; nonetheless, it also has the potential to yield useful information on the improvement in any deficient area of teaching that might be. The way that people show their value priorities might change from society to society (Tarman, 2012; 2016) hence the students in Middle East also have a different perception about SET. A number of studies have emphasized on the need to collect information on teaching quality through students. "Good teaching and good learning are linked through students' experiences of what we do. It follows that we cannot teach better unless we are able to see what we are doing from their point of view" (Ramsden, 2003, quoted in Ali & Ajmi, p. 82).

[1] Ass. Professor of Psychology, University College, Abu Dhabi University, UAE, smitha.dev@adu.ac.ae
[2] Instructor of English & *Freelance Researcher*, *nqayyumhumayun@gmail.com*

There is no question on the objective of such a survey as being inherently honorable and constructive for the overall growth of both the faculty and the institution, but like any other tool of assessment, it is susceptible to an element of error or bias. A study done by a group of faculty members of UAEU concludes that it is not fair to draw comparisons between faculty evaluations while ignoring external factors such as students' GPAs and expected grades, level of the course and its timings, class size and students' gender (Badri, Abdulla, Kamali, & Dodeen, 2006). Many more factors including the age of the faculty member, gender, nationality, appearance, rapport, leniency, course challenge, etc., may also creep in and deflect the true findings. This might be more obvious in diverse and multicultural environments of international universities. Another study on the approach of international students in evaluating their faculty in an Australian university also claims to have received inaccurate and unreliable data through SET, a claim that is supported by empirical evidence (Lama, Arias, Mendoza, & Manahan, 2015).

As mentioned before, the practice of SET is chiefly driven by American system of education; however, most of the times, this procedure is not adapted to the indigenous dynamics of the university using it, especially in terms of culture. For instance, many universities in UAE use it as a routine procedure, but very few are mindful of local differences requiring appropriate modifications. Majority of the students in UAE are of Arabian descent, who are not brought up to judge their teachers, and they might find themselves in a rare position when asked to evaluate them (Sulieman, 2007). Hence, the need of the time is to adapt the SET procedure to draw its real benefit instead of using it as it is, and while doing so, it is also important to understand that this tool had initially been devised for *American students* studying in *American* universities in *America.* Since a number of researchers confirm that "teaching is multidimensional and complex, and therefore, it is difficult to construct a one-fits-all definition of effective teaching" (Al-Hinai, p. 30) see also Adam, 1997; Brown, 1996; Marsh & Dunkin, 1992; North, 1999; Patrick & Smart, 1998), SET tool, the way as we find it, needs to be shaped to suit local subtleties and sensitivity.

It is, therefore, pertinent to investigate all the factors involved in SET ratings and the validity of students' responses in international institutions in the UAE in order to utilize SET tool in the most effective way. In this region, till date, the professors of very few universities, including American University of Sharjah, UAE University, and Higher Colleges of Technology, have done studies on this topic on their respective university's student populations, but the present study is the first of

its kind on the student population of Abu Dhabi University. This study aims to understand the biasing factors in the evaluation of teachers and teaching by the students and in the important decision making by the administrators. Such decisions might be for "improving teaching quality, as well as determining the promotion, contract renewal, and salary increases of teachers pedagogical development and administrative purposes, quality monitoring and control, and making decisions on promotions and tenure" (Rantanen, 2013, p. 224, quoted in Ali & Ajmi, 2013). Current education system is not doing enough to prepare future educators for the demands of non-Eurocentric global education (Kopish, 2016). Thus, the results of this study are expected to benefit the administrators, teachers, and quality control personnel to improvise teaching and teacher evaluation system in a way that brings maximum benefit to higher education (Ali & Ajmi, 2013).

Literature Review

Student Evaluation of Teaching (SET) has become an important instrument in assessing teachers and teaching in the modern world of education, and apparently there is no dearth of research studies that have been done on this topic. Some of the earliest works, which are dated as far back as 1923 were done by psychologist Max Freyd. From then onwards till date, academics have continuously been investigating this process and its implications.

The research in this area has led to different findings, showing very obvious disagreements among researchers on different aspects of this evaluation instrument and leading to an interesting remark by Reckers (1995): ". . . nearly 75 per cent of academics judge student course evaluations as unreliable and imprecise metrics of performance, yet nearly 100 per cent of schools use them, frequently exclusively" (p.33).

More or less, all the literature on SET revolves around three major elements affecting the ratings: factors associated with course, factors associated with teacher, and factors associated with students (Pounder, 2007). On the basis of cognitive dissonance theory, it is argued that poorly performing students give poor ratings to their instructors to protect their self-esteem (Heine & Maddox, n.a.). Moreover, an almost regular pattern of students' liking towards various disciplines and associated ratings has also been observed. According to Cashin, students usually give highest ratings to arts

and humanities courses, whereas social and health sciences are at a medium level, with English language, literature, computing, IT, business, engineering, and physics clustering at the medium to bottom levels (1990). Consequently, the students' passion for the subject of their choice is also reflected in SET scores for that course and its teacher (Marsh and Dunkin, 1992). Elective subjects usually get better ratings since students' liking for the subject constitutes a biasing factor in evaluation.

Among course-related factors, grades and their expectations have a direct influence on SET ratings. Many researchers have agreed on a direct link between the expectation of a high grade with high rating, and the expectation of a low grade with low rating (D'Apollonia and Abrami, 1997; Hudson, 1989; Johnson and Christian, 1990; Mason et al., 1995; Nelson and Lynch, 1984; Perkins et al., 1990; Wilson, 1998; Tata, 1999).

Owing to the fact that "Grades" is a significant variable on SET scores, they have been undoubtedly manipulated by some faculty for their personal benefit. Studies show that faculty have been employing clever grading tactics in terms of easing the assessment procedure by avoiding challenging and contentious teaching material, relaxing grading standards, reducing the amount of teaching and learning material, or spoon-feeding examination content, all eventually leading to grade inflation (Bauer, 1996; Crumbley, 1995; Handlin, 1996; Krautmann and Sander, 1999; Ryan et al., 1980; Sacks, 1996; Schneider, 2013; Simpson and Siguaw, 2000). In brief, university teachers can bargain SET ratings with grades (Hocutt (1987-1988), since they firmly believe that leniency in grades is directly proportional to high SET scores (Martin, 1998; Powell, 1977; Stumpf and Freedman, 1979; Winsor, 1977; Worthington and Wong, 1979; Yunker and Marlin, 1984). Other than the grades themselves, the expectations of grades have also been reported to have a relationship with SET scores, leading the faculty to proactively vouch for good SET scores through grades. There is a:

> . . . kind of mutual back patting taking place where the teacher gives a high grade to the student (this grade not necessarily reflecting any real student attainment) and, in return, the student rewards the teacher with a high teacher rating (Pounder, 2007).

According to Simpson and Siguaw, some university teachers go to the extent of serving snacks on the day of evaluation, praising the class on its performance, or having a "fun activity" before the evaluation (2000).

Teacher's personality is another important factor that impacts SET scores. Teaching is a combination of subject knowledge and teaching skills that help in transferring that subject knowledge to the students; however, such skills should not be confused with the personality traits of teachers that have nothing to do with teaching. Teachers, like other human beings, may have certain characteristics, which despite not having anything to do with their teaching skills, may sway students' liking or disliking for that teacher and act as bias while evaluating them. Feldman (1986), Williams and Ceci (1997), and Cardy and Dobbins also suggest the existence of such a relationship between the teacher's personality and SET scores. As a matter of fact, a study done by Clayson proves that 50 to 80 percent of total variance in SET scores is attributed to the factors related to personality traits of the teachers (1999). Many missed to understand the value of education before they rate their teachers, the goal of 'values education' is the individuals' being sensitive to the events occurring around them and in the world, creating social awareness, honesty and taking responsibility, caring about others, sharing something with others and learning to live together in society (Veugelers ve Kat, 2003; Acun, Demir & Göz, 2010). In the light of such findings, some researchers observe student evaluations as useless numbers that only create competitions among the personalities of faculty members instead of signifying teaching effectiveness (Haskell, 1997; Neath, 1996; Spooren, Mortelmans, & Denekens, 2007, p. 668; Sproule, 2002). Furthermore, some academics have commented that such practice of getting students' feedback ". . . could be a threat to academic freedom" (Haskell, 1997).

Gender is another notable factor involved in SET scores. Matthew Reisz reported in *The Higher Education* that research from France offered evidence that "students appear to rate teachers according to gender stereotypes," with male students giving higher scores to male lecturers. The above report was based on the conclusion of a paper by Anne Boring, a postdoctoral researcher at L'Institut d'Études Politiques de Paris. Her database consisted of 22,665 evaluations by 4,423 first-year undergraduates of 372 different teachers in a single French university. Matthew mentioned in his report that Dr. Boring's analysis suggested that "male students give much higher scores to male teachers in terms of overall satisfaction as well as in all dimensions of teaching." One clear sign of this was that "male students are 30 per cent more likely to rate male teachers' overall satisfaction scores as excellent than when evaluating female teachers" (Boring, 2015).

Research has been done on SET in UAE also. One of the prominent ones in this region is conducted in American University of Sharjah. The study shows that SET ratings are biased without any grain of doubt as ". . . the student's expected grade, teacher's gender, teacher's age, teacher's nationality, teacher's personality, and the students' views of what constitutes "knowledge" are the variables that influence the scores. Other major factors that have strong influence on these scores include origin, gender, language of instruction in high school, and academic status of the students. Compared with the students who had been taught in English or Asian languages, those who had been taught in Arabic in schools were more biased on factors such as the teacher's age, gender, nationality, and personality. This finding might owe its existence to Arab culture in which friendship is correlated with social duties (Al-Issa & Sulieman, 2007).

One more important study was done in UAE University. The study concluded that it would be unfair to compare faculty evaluations without considering the influence of important factors such as "student self-reported GPA and expected grade, course level and timing, class size, and student gender" (Badri, Abdulla, Kamali, & Dodeen, H., p. 51). Their conclusion is in conjunction with the findings of other researchers (Cashin, 1990; Emery et al., 2003; Liaw and Goh, 2003), who argue that using these evaluations for annual faculty appraisals and other important decisions without taking into consideration the above-mentioned variables is a debatable issue. Therefore, "more effort should be directed toward ensuring a more careful interpretation of student ratings in promotion and contract decision processes" (Badri, Abdulla, Kamali, & Dodeen, 2006). Similarly, the age of the teacher (Smith and Kinney, 1992) and the race of the teacher (Smith, 2007) also have adverse effect on students' ratings. Value priorities of teachers and students would be another reason and have an impact on student rating. (Kılınç and others, 2016)

The above documentation is substantial enough for any educationist to reconsider the administering, interpretation, and adaptation of SET tool in other parts of the world than America, and in this study, UAE and the Arab world specifically. As other authors (Becker and Watts, 1999; Boex, 2000; Koh and Tan, 1997; McKeachie, 1997; Tata, 1999) indicated, because of the possible existence of biasing factors in SET, there is a need to supplement it with other measures of gauging teaching effectiveness, and readjust its weight on the overall evaluation of teaching and the teacher.

Objectives

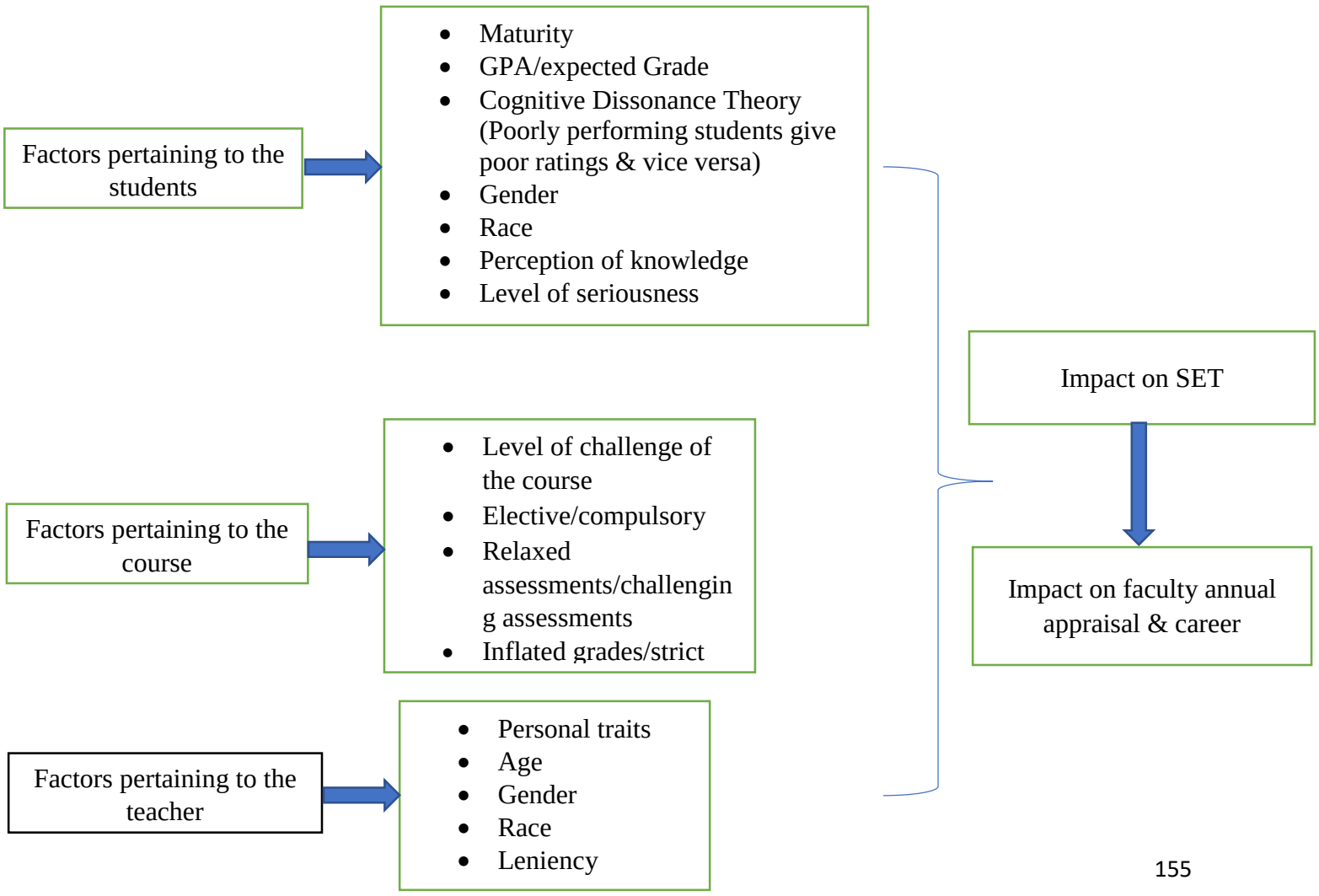
- To understand student's perception towards SET based on age, gender, education, etc.

- To understand various factors affecting student's favorable and unfavorable perception towards SET.

Conceptual framework and hypothesis

- Students have a positive perception towards SET.
- There is a significant difference in the perception of students based on demographic variables such as age, gender, education, nationality, class capacity, and difference in courses.
- Non-instructional factors affect students' perception of SET process.

Figure 1 Conceptual framework



Methods

A quantitative design was used to determine the significant relationship on students' perception and other demographic factors. The researchers conducted the study in 3 different colleges in the university. The survey questions were designed with respect to the objective of the study. The questionnaire was also approved by IRB committee before data collection. Though the sample was selected randomly, the students' willingness to participate in the study was also considered in order to get more accurate results. Since the sample was selected through accessibility, the instrument was administered individually. Few students completed the survey and emailed to the researchers.

Sample

The students that were enrolled in various courses in the university participated in the survey. A total number of 150 students from different colleges like University College, College of Business Administration, and College of Engineering participated. All participants were given equal preference. The questionnaire was distributed among the students after class hours. The instructions were read out, and then the students were requested to look into the given questionnaire. The students were also asked to verify whether they answered all items; the confidentiality of the response was also assured. The demographic data of the sample is shown in the following table.

Table 1: Percentage distribution of the sample according to gender, age, year of study, major, and GPA.

	Count	**Percent**		**Count**	**Percent**
Gender			Year of Study		
Male	56	37.3	2nd year	24	16
Female	94	62.7	1st year	96	64
			Others	30	20
Age			GPA		
17-18	56	37.3	4	10	6.7
19-20	40	26.7	3	73	48.7
Above 20	54	36	2	63	42
			1	4	2.7
Major					
CAS	11	7.3			
COBA	58	38.7			
CECS	57	38			
Others	24	16			

Out of 150 students, more than half (62%) were female students, and only 37.3% were male. Majority of the participants (73.3%) were above 20. 38.7% of the participants were from COBA, 38% were from CECS, and rest of them were from CAS and others. It was noticed that majority (64%) of the students were first years and the others were from 2nd year to final year. Finally, most (90%) of the students' GPA fall in between 2-3, whereas only 6.7% have a GPA of 4.

Tool

A 5-point scale used in this study was a modified version of Student Perception Tool used in American University of Sharjah. The tool was adapted and modified as per our requirement. It consists of 21 items and was used to collect data from the students to understand their perception of SET. It has both positive items (10) and negative items (11). Each item of the scale is related to understand the student's perception towards SET. Maximum score of this scale is 105, and minimum score is 21. Negative items were given negative scoring.

Validity

Validity of the scale was established by correlating the scores of the scale with the scale of Student Perception Scale used in AUS, and validity coefficients were found to be +0.84 and +0.79, respectively. This indicates that the scale is reliable and valid in measuring the perception of students.

Again, the constructed tool was sent to an expert to check whether the contents of the items, individually and as a whole, are relevant to the test. In fact, content validity is the degree in which the test measures are intended in the content area. Content validity was established on the basis of the expert opinions and comments, including those from the IRS Director from Abu Dhabi University.

Reliability

The test retest reliability was established by re - administrating the test and computing reliability coefficient for total test. The reliability coefficient is reported to be +0.92 on a sample of 50 students with the time interval of 2 weeks.

Data analysis and findings

Statistical methods like mean, SD, and chi-squire tests were used to understand the students' perception and also to analyze the association between dependent and independent variables.

Student's perception towards SET

Table 2:

Percentage distribution of the sample according to perception towards SET

Perception towards SET	Count	Percent
Unfavorable	70	46.7
Favorable	80	53.3
Mean ± SD	84.1 ± 10.3	

The overall attitude of the students can be seen in the above table, 46.7% of the students' perception was unfavorable, whereas 53.3% of the students have a favorable perception towards SET.

Association of students' perception towards SET with selected background variables

Table 3:

Comparison of gender, age, year of study, major, and GPA based on perception towards SET.

Gender	Unfavorable		Favorable		χ^2	p
	Count	Percent	Count	Percent		
Male	26	46.4	30	53.6	0	0.964
Female	44	46.8	50	53.2		
Age	Unfavorable		Favorable		χ^2	p
	Count	Percent	Count	Percent		
17 – 18	27	48.2	29	51.8		
19 – 20	17	42.5	23	57.5	0.38	0.827
20 above	26	48.1	28	51.9		
Year of Study	Unfavorable		Favorable		χ^2	p
	Count	Percent	Count	Percent		
2nd year	13	54.2	11	45.8		
1st year	41	42.7	55	57.3	1.68	0.431
Others	16	53.3	14	46.7		
Major	Unfavorable		Favorable		χ^2	p
	Count	Percent	Count	Percent		
CAS	6	54.5	5	45.5	1.86	

COBA	30	51.7	28	48.3		0.602
CECS	25	43.9	32	56.1		
Others	9	37.5	15	62.5		
GPA	Unfavorable		Favorable		χ^2	p
	Count	Percent	Count	Percent		
4	2	20	8	80		
3	31	42.5	42	57.5	5.35	0.069
<3	37	55.2	30	44.8		

It has been observed in Table 3 that the level of perception of students does not significantly vary ($p > 0.05$) based on the variables such as gender, age, year of study, major, or GPA. When we look at the percentage of perception, 53.6% of the male students have favorable perception, whereas 46.4% have unfavorable perception. The percentage of male and female students with their level of perception is almost the same. While observing the variable age, the percentage of perception of students aged 19 and 20 years is higher than the perception of students who are more than 20 years old. Similarly, it has been observed that first-year students are more favorable in their perception when compared with their seniors. 56.1% of the students are from CECS, and these students have a favorable perception towards SET than students from other colleges. Finally, the percentage of favorable perception of the students with higher GPA (4) is better than students with lower GPA (3 and <3).

Table 4:

Percentage of students' perception towards each variable

S. No.	Statement	1-SD	2- D	3-NS	4-A	5-SA
		%	%	%	%	%
1	By evaluating my professor, I am actually helping them improve their teaching effectiveness	34.67	34.00	24.00	6.00	1.33
2	Professors change their teaching methods as a result of student evaluations	15.33	38.67	30.00	12.67	3.33
3	The Course Evaluation Form is adequate enough to evaluate my professors	14.67	41.33	30.00	9.33	4.67
4	ADU students should take faculty evaluations seriously	44.67	35.33	15.33	3.33	1.33

5	I fill out all the questions including recommendations and suggestions	25.33	38.00	27.33	7.33	2.00
6	I read and understand each statement before I rate it	46.00	32.00	19.33	2.00	0.67
7	My rating of my professors is affected by my expected grade in the course	9.33	18.67	26.67	20.67	24.67
8	I am comfortable taking courses with male professors	39.33	37.33	13.33	6.67	3.33
9	I am comfortable taking courses with female professors	35.33	38.00	18.67	5.33	2.67
10	I prefer taking courses with young and enthusiastic professors	22.00	22.00	44.00	9.33	2.67
11	I prefer taking courses with older and experienced professors	18.67	32.67	34.00	8.00	6.67
12	I am more comfortable with Arabic-speaking professors	20.00	24.67	28.67	11.33	15.33
13	I am more comfortable with non-Arabic-speaking professors	18.00	32.67	30.67	12.00	6.67
14	When evaluating my professors, I usually pay more attention to their personality (i.e., friendless, leniency, looks, dress, etc.)	17.33	37.33	24.00	13.33	8.00
15	If I have a good relationship with my professor, I will rank him/her high on teaching effectiveness	12.67	32.00	23.33	21.33	10.67
16	If I ask my professor a question that is related to the subject being taught and my professor responds by saying "I am not really sure, but I will check on that and get back to you," I will still not consider him or her knowledgeable	10.67	15.33	25.33	27.33	21.33
17	If I have a spoiled relationship with my professor, I will rank him or her low on teaching effectiveness	8.00	12.67	32.67	24.00	22.67
18	My responses are always honest while evaluating my professors	41.33	40.00	14.67	2.00	2.00
19	I always understand the seriousness of the SET evaluation process	27.33	46.67	17.33	8.00	0.67
20	The class timings of our lecturer affect my evaluation ratings	8.00	20.67	30.67	18.67	22.00
21	The class capacity (more or less number of students in the class) affects my evaluation rating on professors	10.67	17.33	33.33	19.33	19.33
22	The qualification of my professor affects my evaluation rating on my professor	12.67	28.00	32.67	17.33	9.33
23	I enjoy graduate courses more than the UC courses	36.67	23.33	33.33	6.00	0.67
24	ADU should continue having students evaluate their professors	52.67	28.00	13.33	5.33	0.67

Since we could not observe any statistically significant association between the variables, we did a simple percentage analysis to understand more about students' perception. It was observed that

67% of the students strongly disagree that by evaluating the professor, they are helping them to improve their teaching, whereas only 7.33% of the students have a positive perception, they believe that by evaluating their professor, they are helping them to improve their teaching effectiveness It was also noted that the students do not take the task of filling out the questionnaire seriously. That is evident in the statements 4, 5, 6, 18, and 19 of the above table, which, to be more specific, mean that they complete the evaluation form without even reading, understanding, and filling it out completely. The above also shows that they are not honest in their responses.

Discussion and Conclusion

In general, the findings of the present study show that ADU students have a positive perception towards SET evaluation and the stated hypothesis is accepted. However, the research could not observe any statistically significant difference between students' perception towards SET and other demographic variables such as age, gender, year of study, major, and GPA.

The findings show that, overall, students do not take SET process seriously. They do not believe that the teaching process will improve or the teachers will change their teaching through their evaluations. They admit to not filling out the form completely and not considering it adequate. Their response to statement 6 in Table 4 shows that they do not even take the trouble to read and understand each question in the SET sheet before responding to it. A spoiled relationship with the teacher also influences their evaluation of that teacher. The most serious finding that needs significant amount of attention from educationists is the students' frank acceptance of their dishonesty in the evaluation process: 41.33% *strongly disagree* and 40.00% *disagree* with the statement that they are honest in evaluating their professors. Additionally, their seriousness towards this process of evaluation of their teachers is checked twice, both directly and indirectly, and the responses through both the questions confirm that their approach towards this procedure is the least bit serious. However, gender, age, or race do not show any significance in their evaluation of teachers.

The above results are in alliance with several other research papers. They visibly prove that SET scores are not reliable, and there is a strong need of supplementing this evaluation system with other measures and assessments to get a more dependable picture of a teacher and his or her teaching.

Suggestions for Improvement in the Evaluation System

It is impossible to create a fool proof system of a teacher and teaching evaluation; however, in order to ensure better evaluation, any such system should be continuously refined so that a picture closest to the real class room dynamics can be derived.

The results of this research lead to a few recommendations. To begin with, faculty members must have a strong hand in the development of such an assessment tool. Each college has to take responsibility of developing its own evaluation method and evaluation criterion basing upon the class environment and the emotional maturity of the student population.

Almost all the universities incorporate a three-pronged plan for assessing the competence of their faculty: student evaluation, peer evaluation, and self-evaluation. Although this plan appears to be a clever strategy, which it also is, it is not perfect and is liable to bias. That bias can be neutralized to some extent if the weights of all the components of that evaluation plan are reduced. Secondly, both very high and very low evaluation scores should be investigated. Further, students should be educated on the nature and consequences of SET results and should also be given short practice prior to administering the original process to train them for the real task. Also, in second language learning environments, the form should include translation of the questionnaire. Another way of gathering important and somewhat real information about teaching standard of a teacher is through spot checks. These checks could be done through visiting classes briefly, reviewing supplementary course materials, reviewing question papers, and talking to students. The question papers picked for spot checks should be any assessment components other than the final exam paper, for which teachers usually take extra care to avoid issues.

This revision in SET procedure and incorporation of additional strategies for evaluating the real level of a teacher's teaching are nothing less than crucial, not only because the above mentioned unethical teaching practices to generate high scores in SET promote dishonest teachers getting high scores and honest teachers getting low scores, but also because through an incorrect reflection of a teacher's performance, a truly worthy teacher can miss acknowledgement and suffer grave consequences on his or her career.

References

Acun, İ, Demir, M. & Göz, L. (2010). Öğretmen adaylarının vatandaşlık yeterlilikleri ile eleştirel düşünme becerileri arasındaki ilişki, *Journal of Social Studies Education Research, 1*(1), 107-123.

Al-Hinai, N. S. (2011). Effective college teaching and student ratings of teachers: What students think, what faculty believe, and what actual ratings show: Implications for policy and practice teaching quality assurance and control in higher education in Oman. *Unpublished doctoral dissertation*, Durham University, UK.

Ali, H. I., & Ajmi, A. A. (2013). Exploring Non-Instructional Factors in Student Evaluations. *Higher Education Studies*, 3(5), 1-13.

Al-Issa, A., & Sulieman, H. (2007). Student evaluations of teaching: perceptions and biasing factors. *Quality Assurance in Education*, 15(3), 302-317.

Arreola, R. A. (2007). Developing a comprehensive faculty evaluation system: A guide to designing, building, and operating large-scale faculty evaluation systems. San Francisco, CA: Anker Publishing.

Badri, M. A., Abdulla, M., Kamali, M. A., & Dodeen, H. (2006). Identifying potential biasing variables in student evaluation of teaching in a newly accredited business program in the UAE. Emerald Insight: *International Journal of Educational Management*, 20 (1), 43-59. https://doi.org/10.1108/09513540610639585

Bauer, H.H. (1996). “The new generations: students who don’t study.” Paper presented at The Technological Society at Risk Symposium, Orlando, FL (10 September 1996)

Becker, William E. and Michael Watts. (1999). "How Departments of Economics Evaluate Teaching." *American Economic Review*, 89(2), 344-49.

Boring, A. (2015, August 14). Female lecturers ‘suffer from gender biases’ in student ratings. Retrieved from The World University Rankings. https://www.timeshighereducation.com/news/female-lecturers-suffer-gender-biases-student-ratings

Cardy, R.L. and Dobbins, G.H. (1986). “Affect and appraisal accuracy: liking as an integral dimension in evaluating performance.” *Journal of Applied Psychology*, 71(4), 672-8.

Cashin, W. E. (1990). Students do rate different academic fields differently. In M. Theall, & J. Franklin (Eds.), Student ratings of instruction: Issues for improving practice: *New Directions for Teaching and Learning*, 43, 1990, 113-121.

L. F. Jameson Boex. (2000). Attributes of Effective Economics Instructors: An Analysis of Student Evaluations. *The Journal of Economic Education*, 31(3), 211-227. doi:10.2307/1183092

Clayson, D.E. (1999), "Students' evaluation of teaching effectiveness: some implications of stability." *Journal of Marketing Education, 21*(3), 68-75.

Crumbley, D.L. (1995), "The dysfunctional atmosphere of higher education: games professors play." *Accounting Perspectives*, 1(1), 27-33.

D'Apollonia, S. and Abrami, P.C. (1997), "Navigating student ratings of instruction." *American Psychologist, 52* (11), 1198-208.

Emery, C. R., et al. (2003). Return to academic standards: A critique of student evaluations of teaching effectiveness. *Quality assurance in Education, 11*(1), 37-46.

Feldman, K.A. (1986). "The perceived instructional effectiveness of college teachers as related to their personality and attitudinal characteristics: a review and synthesis." *Research in Higher Education, 24(2),* 139-213.

Handlin, O. (1996). "A career at Harvard." *American Scholar, 65* (5), 47-58.

Haskell, R. E. (1997). Academic freedom, tenure, and student evaluation of faculty: Galloping polls in the 21st century. *Education policy Analysis Archives, 5*(6), 1-32.

Heine, P. and Maddox, S. (2009). Student perceptions of the faculty evaluation process: An exploratory study of gender and class differences. *Research in Higher Education Journal*, 3(1), 1-10.

Kilinç, E., Kilinç, S., Kaya, M., Başer, E., Türküresin, H., & Kesten, A. (2016). Teachers' attitudes toward the use of technology in social studies teaching. *Research in Social Sciences and Technology, 1*(1), 59-76.

Hocutt, M.O. (1987-1988), "De-grading student evaluations: what's wrong with student polls of teaching", Academic Questions, Winter, 55-64.

Masood A. Badri Mohamed Abdulla Mohammed A. Kamali Hamzeh Dodeen, (2006) "Instructional collaboration: creating the learning environment." in Fradd, S.and Wismantel, P. (Eds), Meeting the Needs of Culturally and Linguistically Different

Students: (Review from A Handbook for Educators, by Hudson) International Journal of Educational Management, 20(1), 43 – 59. http://dx.doi.org/10.1108/09513540610639585

Johnson, R.L. and Christian, V.K. (1990). *Perceptual and Motor Skills, 70*, 479-82.

Koh, C.H. and Tan, T.M. (1997). "Empirical investigation of the factors affecting SET results." *International Journal of Educational Management, 11*(4), 170-8.

Kopish, M. A. (2016). Preparing globally competent teacher candidates through cross-cultural Experiential learning. *Journal of Social Studies Education Research, 7*(2), 75-108

Krautmann, A. C., & Sander, W. (1999). Grades and student evaluations of teachers. *Economics of Education Review, 18*(1), 59-63.

Lama, T., Arias, P., Mendoza, K., & Manahan, J. (2015). Student Evaluation of Teaching surveys: do students provide accurate and reliable information? *E-Journal of Social & Behavioral Research in Business. 6*(1), 38.

Marsh, H. and Dunkin, M. (1992), Students' Evaluations of University Teaching: *Handbook on Theory and Research, 8* (pp. 143-234). Agathon Press, New York, NY.

Martin, J.R. (1998). "Evaluating faculty based on student opinions: problems, implications and recommendations from Deming's theory of management perspective." *Issues in Accounting Education, 13*(4), 1079-94.

Mason, P., Steagall, J. and Fabritius, M. (1995). "Student evaluation of faculty: a new procedure for aggregate measures of performance." *Economics of Education Review, 14*(4), 403-16.

McKeachie, W.J. (1997). "Student ratings: the validity of use." *American Psychologist, 52* (11), 1218-1225. http://dx.doi.org/10.1037/0003-066X.52.11.1218

Neath, I. (1996). How to improve your teaching evaluations without improving your teaching. Psychological Reports, 78, 1363-1372. doi:10.2466/pr0.1996.78.3c.1363

Nelson, J.P. and Lynch, K.A. (1984). "Grade inflation, real income, simultaneity, and teaching evaluations." *Journal of Economic Education, 15*(1), 21-39.

Newton, J. (1988). "Using student evaluation of teaching in administrative control: the validity problem." *Journal of Accounting Education, 6*(1), 1-14. https://doi.org/10.1016/0748-5751(88)90033-4

Perkins, D., Gueri, D. and Schleh, J. (1990). "Effects of grading standards information, assigned

grade, and grade discrepancies on student evaluations." *Psychological Reports, 66*(2), 35-42.

James S. Pounder, (2007) "Is student evaluation of teaching worthwhile?: An analytical framework for answering the question." *Quality Assurance in Education, 15*(2), 178-191, https://doi.org/10.1108/09684880710748938

Powell, R.W. (1977). "Grades, learning, and student evaluation of instruction." *Research in Higher Education, 7*(3), 193-205.

Reckers, P.M.J. (1995). "Know thy customer", in Baril, C.P. (Ed.), Change in Accounting Education: A Research Blueprint, Federation of Schools of Accountancy, St Louis, MO.

Ryan, J.I., Anderson, J.A. and Birchler, A.B. (1980).." *Research in Higher Education, 12* (4), 317-33.

Schneider, G. (2013). Student evaluations, grade inflation and pluralistic teaching: Moving from customer satisfaction to student learning and critical thinking. *Forum for Social Economics Journal, 42*(1), 122-135.

Sacks, P. (1996), Generation X Goes to College. *ERIC Digest.* Retrieved from ERIC database. ED399906 Open Court, Chicago, IL.

Seldin, P. (1989). *American Association for Higher Education Bulletin, 41*(7) 3-7.

Simpson, P.M. and Siguaw, J.A. (2000). *Journal of Marketing Education, 22*(3), 199-213.

Smith, B. P. (2007). Student Ratings of Teaching Effectiveness: An Analysis of End-of-Course Faculty Evaluations. College Student Journal, The University of Georgia, *41*(4), 788+.

Spooren, P., Mortelmans, D., & Denekens, J. (2007). Student evaluation of teaching quality in higher education: Development of an instrument based on 10 Likert-scales. Assessment & *Evaluation in Higher Education Journal, 32*(6), 667-679.

Sproule, R. (2002). The underdetermination of instructor performance by data from student evaluation of teaching. Journal of Economics of Education Review, *21*(3), 287-294.

Stratton, W.O. (1990), "A model for the assessment of student evaluations of teaching, and the professional development of faculty", *The Accounting Educators' Journal*, 17(2), 77-101.

Stumpf, S. A., & Freedman, R. D. (1979). Expected grade covariation with student ratings of Instruction: Individual versus class effects. *Journal of Educational Psychology, 71*(3), 293-302. http://dx.doi.org/10.1037/0022-0663.71.3.293

Tata, J. (1999). “Grade distributions, grading procedures, and students’ evaluations of instructors: a justice perspective.” *The Journal of Psychology,133* (3), 263-71.

Tarman, B., & Acun, I. (2010). Social studies education and a new social studies movement. *Journal of Social Studies Education Research, 1*(1), 1-16.

Wilson, R. (1998). “New research casts doubt on value of student evaluations of professors.” *The Chronicle of Higher Education, 44*(19), A12-A14.

Winsor, J.L. (1977). A's, B's, but not C's?: A Comment” *Contemporary Education, 48*(2) 82-4.

Worthington, A.G. and Wong, P.T.P. (1979). Evaluations of an instructor.” *Journal of Educational Psychology, 71*(6), 764-775. http://dx.doi.org/10.1037/0022-0663.71.6.764

Yunker, J.A. and Marlin, J.W. (1984), “Performance evaluation of college and university faculty: an economic perspective”, *Educational Administration Quarterly*, 15(2), 9-37.

Zhao, J., & Gallant, D. (2012). Student evaluation of instruction in higher education: Exploring issues of validity and reliability. *Assessment & Evaluation in Higher Education, 37*(2), 227-235.