

# STUDENT EVALUATION OF TEACHING, FACULTY PERCEPTIONS: CASE OF AN ASIAN BUSINESS SCHOOL

Valentine Matthews, Lancaster University

## ABSTRACT

*This research sought to further knowledge on faculty perceptions of student evaluation of teaching. Faculty perceptions of student evaluations of teaching at institutional, academic programs and course levels have been published for at least a century. The focus of this study is at program and institutional level of a Business School within Asia to query the outcomes and understand faculty members' perceptions of this tool. It also elicits the challenges faced by these evaluations in the context of the case institution and its relationship with the faculty's self-evaluation of their own teaching.*

*Mixed methods were employed using primary and secondary survey data. The study collects students and faculty ratings and comments are requested to evaluate the effectiveness of teaching in the HEI. Medium and positive correlation is found between course and instruction ratings. Faculty are found to be satisfied with the use of the faculty evaluation of teaching above other alternatives although a combination of multiple tools is deemed acceptable by faculty. Faculty self-evaluation of teaching and student evaluation of teaching is found to be positively correlated. The research exposes similarities between the outcomes of student evaluation of teaching and the faculty self-evaluation of teaching. It also suggests that student and faculty ratings are clustered at the upper end of the rating scale. The implication is that student evaluation of teaching is preferred by faculty of the case higher education institution over faculty self-evaluation of teaching, peer evaluations, and consultant evaluations as a tool for evaluation of the quality of teaching and learning. The findings demonstrate that results buttress existing research and therefore could be further investigated.*

**Keywords:** Student Evaluation of Teaching, Faculty Evaluation of Teaching, Effectiveness, Quality.

## INTRODUCTION

Student Evaluation of Teaching (SET) surveys are commonly employed for evaluating the quality of teaching and learning (Gustad, 1961; McKeachie, 1997; Charteris & Smardon, 2019) in Higher Education Institutions (HEIs). Its scope of use is dependent on the quality context (Hou et al., 2022). Its widespread adoption is attributable to the ease of gathering, presenting and interpreting the collected responses (Penny, 2003 p.2). Outcomes of the SETs theoretically have significant implications for quality of teaching (Johnson, 2000). SETs are also used in performance evaluation (PE) of faculty (Seldin, 1993; Nowell et al., 2009) and queries student voice for teaching, learning and decision making (Lyde et al., 2016). Tools such as Faculty Self-Evaluation of Teaching (FSET), peer evaluations (PE) and independent consultant evaluations, supervisor evaluations (Borgen and Davis, 1974) are also independently adopted or for complementing SETs for PE (Benton and Young, 2018). FSET describes faculty self-perception of 'attributes of their own teaching' (Lyde et al., 2016).

The design of the content of the SET is not universal and is specific to individual HEI. To extract pertinent information to inform improvements, implementation and

interpretation of SET is done at course, program or at whole HEI levels (Schultz and Latif, 2006; Fricks et al., 2007; Arbula Blecich & Zaninović, 2019). This suggests that conducting the SET at the micro-social domain or at the level of the faculty extracts student perception of the course. However, for the purposes of programs and institutional monitoring, the analysis is at the macro-level. Evaluating teaching at institutional level produces overall HEI information but suffers from significant loss of useful inference for training and improvement at unit level. The outcomes of the SETs are most pertinently deployed at the level of the faculty for improvement of the teaching and for addressing other student concerns. Studies on the relationships between outcomes of the SET and FSET may inform the use of these tools for decisions on faculty performance. Are there similarities between the SET and FSET ratings by faculty and students?

## LITERATURE REVIEW

### Effective teaching

Effective teaching is subjectively defined and known by observers when seen (Cashin, 2003; Benton and Cashin, 2012; Miron & Mevorach, 2014). Good et al., (2009) viewed effective teaching as ‘the ability to improve student achievement’ thus ascribing a cause-and-effect relationship between teaching and student performance. Effective and excellent teachers may possess attributes such as interesting teaching style, accessible personal traits, clear teaching techniques (Alhijia, 2017; Pan et al., 2009) teacher-student relationship, engagement and real-world experience (Pan et al., 2009; Keeley et al., 2006; Liu et al., 2016; Keeley et al., 2012; Liu et al., 2015). Institutional accountability thus requires HEIs to implement adequate measures for assuring stakeholders that the quality of teaching is being improved (Kember et al., 2002).

### Student Evaluation of Teaching (SET)

SETs are the foremost tools for evaluating teaching or teaching effectiveness (Gustad, 1961; McKeachie, 1997; Spooren et al., 2013). In a Turkish study, SET data collected from students was seen as valid for evaluating faculty performance (ÜSTÜNLÜOĞLU & Seda, 2012). Seldin (1993) in a US study observed that the number of HEIs’ implementing the SET had increased from 28% in 1973, to 86% in 1993. Faculty development (Park et al., 2020; Theall, 2017) are also informed by the outcomes which also serve as evidence of institutional teaching effectiveness (Denson et al., 2010). The SET is fundamental in faculty evaluation or sometimes used as the sole means for faculty evaluation (Seldin, 1993). Richardson (2005) noted that ‘HEIs invest more efforts on student satisfaction with teaching and programs.’ SETs are also implemented in courses as a quality assurance tool to survey student satisfaction with the instruction in courses (Johnson, 2000).

SET outcomes and ‘*student characteristics, physical environment, grading leniency and instructor*’ (Arbula Blecich & Zaninović, 2019; Spooren 2010) and learning (Marsh, 1987) have been reported to be highly correlated. On the contrary, some studies found SET outcomes to be independent of student gender and grades (Fricks et al., 2007) and field of study (Spooren and Van Loon, 2012). Basow and Silberg (1987) and Boring, et al. (2016) found student bias against female faculty in multiple countries. According to Basow and Silberg (1987), both male and female students rated the female faculty members lower than the male faculty. Course discipline, sexual orientation and race of the faculty are also seen

as biasing factors of SET outcomes (Spooren et al., 2013). Such biases and other factors not tested by the SET could likely skew the SET outcomes. There is not a universally tested and agreed SET as their contents of the questions are often context specific. Harvey in Penny (2003) looking at faculty personality influence saw the SET as the “*happy form*”. This suggests SET outcomes may be biased by faculty personality (Shah et al., 2020).

SET also informs decisions on faculty salary increases and retention (Benton & Young, 2018; Kulik, 2001). Faculty were found to disagree with outcomes of SETs for promotion or salary increases (Iyamu & Aduwa, 2005). With so many biases, the outcomes of the SETs are often disputed by faculty. SET outcomes have sometimes shown that faculty with best reviews were not necessarily the most effective (Surratt & Desselle, 2007). Students’ lack of training in implementing course materials is also seen by faculty as an invalidating factor of the SET outcomes (Mohammad, 2001; Gaillard et al., 2006) however, SET queries wider experiences than the curriculum.

SETs are commonly implemented by Middle East and Asian HEIs as instruments for measuring faculty effectiveness. This is reflected in Jordan, Omani and Chinese contexts (Alsmadi, 2005; Yin et al., 2014; Al-Maamari 2015; Yin et al., 2016). Faculty gender and class size were identified in an Omani study as weak predictors of SET outcomes. Faculty have not unanimously received the SET (Iyam & Aduwa-Oglebaen, 2005; Mwachingambi & Wadesango, 2011) while other publications in Africa and Asia show that faculty accept the SET only for formative purposes (Idaka et al., 2006; Inko-Tariah, 2013). These findings justify the use, validity and reliability of SET as measure of quality of teaching.

Students are good sources of informed data on the quality of teaching assessments (Cashin, 1989; Chism, 1999). Most studies implemented mixed methods or strictly quantitative research method (Harun et al., 2011); Iyamu & Aduwa, 2005; Chan et al., 2014) in researching SETs. Mixed method for eliciting student perceptions of aspects of the teaching, learning, and curriculum may provide clearer insights into performance of individual faculty. Qualitative student comments allow the students to provide unguided creative or innovative ideas beyond the HEI’s planned ideas. Whatever the adopted method taken by the SETs, literature shows that they are employed by HEIs as measures of faculty performance using the customer feedback in judging the effectiveness of institutional teaching and learning. Students’ award of similar ratings across multiple evaluation questions (Spooren, 2010) have also been criticised by faculty. This is seen as a bias in the survey tool or in the completion of the surveys by the students. The practice of implementing the surveys by some HEIs before the course assessments is done to minimise the impact of grades on the SET outcomes. However, this approach could be biased depending on perceptions of the course assessments as a good measure of the teaching.

### **Faculty Perceptions of SET**

SETs are often criticised by faculty over its use for evaluating the quality of teaching (Spooren et al, 2013; Boring et al., 2016; Hornstein, 2020) and its insufficiency in addressing the meaning of “effective teaching”. Faculty also raise concerns over its validity, reliability of outcomes, (Boring et al., 2016, Hornstein, 2020 and Spooren et al, 2013) students, faculty and course characteristics (Spencer & Schmelkin, 2002; Wachtel, 1998). These challenges of the reliability, validity and concerns with SETs has not impeded its adoption by HEIs. Faculty perceptions of SEIs, its purpose, validity, usefulness, and consequences remain mixed (Wachtel, 1998; Surgenor, 2013). The faculty perceptions of the SET, its challenges and its sufficiency in evaluating faculty performance is desired. The evaluations, development, salary and promotion of the faculty hinges largely on the

outcomes of the SET hence it is necessary to understand their perceptions of the tool and other evaluative tools such as FSET, peer evaluations, and consultant evaluations.

### **FSET and other evaluative tools**

There are indications that faculty are satisfied with the use of SETs as a tool for formative feedback on the instruction and instructor (Balam & Shannon, 2010; Kulik, 2001; Spooren, 2013; Shah et al., 2020). This suggests a likelihood of a relationship between the outcomes of SETs and FSETs. High correlation between the SET and FSET suggests that there is a redundancy of combining these tools for assessing faculty. Feldman (1989) found correlation between SET and FSET to be low while medium correlation below 0.5 was found by other studies (Marsh et al., 1979; Marsh, 1987; Marsh & Dunkin, 1992; Beran & Violato, 2005). These correlations scores dispute faculty concerns over the validity of SETs. The effectiveness of FSET is likely a challenge due to its subjectivity which again raises questions of its use for formative evaluations and accountability tool (MacBeath, 2003) and for faculty PE (Marsh, 1984; Cashin, 1990; McKeachie, 1997; Atwater et al., 2002; Lyde et al., 2016; Moreno et al., 2021).

Academic course is seen as an explanatory variable for the outcome of the FSET (Idaka et al., 2006; Jackson, 1998; Harun et al., 2011; Al-Maamari, 2015), especially where the courses are from different schools or colleges like engineering and arts. This would apply at macro-levels and is unlikely to apply at micro and meso-levels. Academic course could impact FSET conducted at institutional level with multiple fields and colleges. However, faculty training, knowledge of the curriculum and teaching approaches would suggest that FSET outcomes could be more objective than the SET outcomes. There are risks associated with the investment of trust in faculty to objectively self-assess. HEI's tend to adopt a combination of feedback sources to inform PE (Entrekin & Chung, 2001; Atwater et al., 2002). Use of these combinations of evaluative tools may be challenged where strong correlations exist between some of these tools. Strong correlations between evaluative tools indicates a redundancy that needs to be addressed by using only one of the two related tools.

Faculty oppose the use of peer evaluations for summative assessments however, it is favoured for formative developmental purposes (Berk, 2005). Criticisms of faculty is that peers with insufficient experience of quality teaching may produce biased reports (Chism, 1999 p. 75). Use of external independent consultants on the other hand is impaired by their likely low knowledge of the local context to arrive at appropriate evaluation of teaching. It is important to understand faculty members' order of preference of these tools and their justifications for the ranking of the tools. The research seeks to understand the outcomes of the SET and its use for evaluating the effectiveness of the faculty's teaching in a Business School context in the Middle East. It also seeks to elicit faculty perception of the case HEI's SET, faculty perceived challenges of the SET, and if the SET and FSET outcomes correlate.

## **METHODOLOGY**

### **Research Design**

Mixed methods approach is adopted in studying the research questions (Chikazinga, 2018). Mixed quantitative and qualitative approach is largely more suitable for extracting meso- and macro-level information (Chen et al., 2014; Lyde, 2016) in HEIs. Micro-level study of individual faculty performance and perceptions are more appropriately studied

using interviews. Macro-level study requires the use of larger data sets (Sulong, 2018, Chen et al., 2014 and Ginns et al., 2007) creating opportunities for making inferences. This study is a meso-level study and employs primary qualitative and a combined primary and secondary quantitative data retrieved by the HEI and in a faculty perceptions survey. To minimise the number of surveys completed by the recipients, the FSET is included in the faculty perceptions survey. The primary and secondary data are perceptions of participants' lived experiences. These perceptions are subjective constructs of the social reality of the faculty and students in this context (Uwe Flick, 2009; Shah et al., 2020). The adopted ontological approach sees individual faculty as having their views of the SET. A constructionist paradigm is implemented in arriving at an institutional perspective. The research also assumes an epistemological stance that knowledge needs to be interpreted to discover the underlining meanings in the data.

## Research Context

The case HEI is particularly situated in the Middle East region. The region is actively improving its higher education sector in response to global competitions for skills development. The HEI is a small Business School, established about 15 years prior and delivers graduate level business and leadership related academic programs. It has similar characteristics as the research context of Basow and Silberg (1987) however, they are situationally distinct. The HEI is characterised by having at most 20 academic faculty members at any time. The academic programs are graduate degrees with a length of 1.5 years. The maximum number of students in the HEI in an academic year is 400. The HEI is also strongly characterised by the high number of faculty with an average of 8 years' experience in HEIs.

## Data Collection Tools

The case HEI's SET, FSET and faculty perceptions questionnaire have quantitative ratings and qualitative comments sections. The faculty perception questionnaire has 32 questions clustered into: perception of the implemented SET, perceptions of alternatives to SET such as peer review of teaching (Henderson et al., 2014; Douglas and Douglas, 2006), consultant evaluation, FSET, factors challenging use of SET for PE, disadvantage of the implemented SET, comments, and a section on the demographics. In the USA, peer review is implemented as a mechanism for feedback on aspects of teaching requiring further development (Osburne & Purkey, 1995). The questionnaire was piloted on 6 faculty members and the Cronbach alpha from this pilot is 0.714. Cronbach Alpha reliability coefficient close to 1 validates for the faculty perception of SET questionnaire (Iyam and Aduwa-Oglebaen, 2005; Harun et al., 2011; Chan et al., 2014; Yin et al., 2016). The implemented questionnaire has 1 question ranking statement, 25 Likert ratings, and 6 comments. The questionnaire is adapted from Deaker et al. (2010) which has a mix of Likert scale, Yes/No, and "open-ended comments" questions. Likert-type scales are ordinal scales of series of qualitative statements that respondents use to indicate and convey their opinion or perceptions of evaluative questions (Vogt, 1999). The Likert scale adopted in this study range from strongly disagree (1), disagree (2), neither agree nor disagree (3), agree (4) and strongly agree (5).

The SET questions reflect the aspects most relevant to the context, educational vision, policy, and in accordance with its perception of teacher quality (Penny, 2003). The HEI's SET has three sections: Part A (course), Part B (instructor) and Part C (comments). Sections A and B together have a Cronbach Alpha of 0.926 and contain 5 questions each

with a 5-point Likert scale response. Section C as aptly titled is a written-in comment section. The SET is an anonymised survey implemented over 15min by the institution during the course and prior to the summative assessment to nullify possible bias of outcomes of the assessments.

## Sampling

Purposive sampling approach was adopted because the sampling population of faculty is 20. However, participation in the study by 11 faculties was voluntary. Table 1 describes the characteristics of the participants from all 4 academic programs. The faculty mean experience at the HEI is 3.11 years and their mean experience in HEIs generally is 13.28 years. This implies that the faculty possess significant experience in higher education generally in the HEI and that they have experience with implementing SETs and/or FSETs.

Faculty	Academic Program	Gender	Academic Rank	Case HEI Exp.(yrs)	HE Exp. (yrs)
X1	MSBA	F	Asst. Prof.	1.20	5.10
X2	MBA	M	Asst. Prof.	5.50	15.00
X3	MSLOD	M	Asst. Prof.	3.50	12.50
X4	MSBA	F	Asso. Prof.	1.30	7.00
X5	MSQBA	M	Asso. Prof.	5.30	6.50
X6	MBA	M	Asso. Prof.	3.50	25.60
X7	MBA	M	Asso. Prof.	5.80	10.50
X8	MSLOD	M	Asso. Prof.	2.90	22.50
X9	MSLOD	F	Asso. Prof.	0.90	25.30
X10	MSQBA	M	Asso. Prof.	3.60	12.60
X11	MBA	M	Asst. Prof.	0.70	2.70
Mean				3.11 (1.81)	13.28 (8.11)

## Data Analysis

The correlation or strength of relationship between the SET and FSET at the institutional level is done by implementing a linear regression of outcomes of the same questions. The HEI delivers 4 academic programs all within one college. Hence the institutional perspectives could be researched from the single college. Researcher interpretation is imposed on both the written-in students and faculty comments from the SET and faculty perception of the SET qualitative comments, coded and categorized. Content analysis is implemented to objectively, systematically and quantitatively depict the content of purposively collated communication of suggestions by students for the improvement of the instruction (Berelson, 1952). These student comments provide suggestions on SET implementation challenges, improvement of the instruction. Categorization of the codes into smaller collectives of meanings is done (Weber 1990: p.15). The numerical data generated from the categorization of the comments are hence analysed using descriptive statistics.

## RESULTS

### SET in evaluating teaching effectiveness

Table 2 presents the outcomes from the analysis of the secondary SETs implemented midway through delivery of 11 courses by faculty members. Class size in the courses is  $12 < x < 40$ , with a mean of 24.46, and the sum of students in the 11 courses is 296. The overall student participation rate in SET implemented on the randomly selected courses is 68%. The mean student ratings of the effectiveness of the course materials and instruction are between 4.45 and 4.54 respectively. These suggest that the students on the average agree or strongly agree that the faculty's course materials and instruction are effective. Outcomes of faculty X10 is an outlier due to its low mean rating of  $3.78 \pm 0.18$  for the course materials and  $3.62 \pm 0.90$  for the instruction. The students rated faculty X10's course and instruction as less than effective.

Outcomes of the analysis of the HEI's aspects of course materials and instruction (see table 2) are presented in table. Apart from the ratings for faculty X10, the mean ratings of the course evaluation question A1 ... A5 are generally lower than the scores of the instruction evaluation questions (B1, ... B5). The students generally view the course materials and the instructions as at least effective.

Aspects	Course Materials					Mean	Aspects of Instructions					
	A1	A2	A3	A4	A5		B1	B2	B3	B4	B5	B6
Mean	4.54	4.47	4.41	4.40	4.41	4.45	4.56	4.54	4.57	4.53	4.51	4.54
SD	0.27	0.30	0.22	0.30	0.27	0.25	0.36	0.33	0.40	0.37	0.39	0.37

### Relationships between SET course and instruction ratings

By implementing a linear regression analysis between the SET means of course materials and instruction by faculty was used to determine if there is a correlation between both variables. Table 3 provides the outcomes of the implemented regression analysis between each faculty's student rating of the course materials and the instruction. This sought to establish the degree of relationship between the faculties's rating of the course materials and the instruction. With  $r = 0.95$ ,  $\alpha = 0.05$  and the p-value 0.00001, the findings show that the ratings of the course materials and instruction are strongly correlated. The finding agrees with Spooen (2010) that the course ratings and instruction ratings are similar however, the findings deviate regarding the strength of the relationship. This tends to ensure that students continue to show gratitude for the teaching without being influenced by the course assessment grades as noted by Zimmerman (1998) and Corno (2001). It could be argued that the SET as implemented assesses students' opinions within the "*happiest period*" of the course because the final course grades which would likely impact their opinions have not been received. Students would likely blame poor performance on the instruction.

	Coefficients	Standard Errors	T stats	P value
Intercept	1.464	0.328	4.466	0.00156
Instruction	0.657	0.072	9.126	0.00001

### Findings in the SET comments

The SET question solicited student responses on how the instruction and materials could be improved. The outcomes of students' response this question. Five categories were generated with seventeen codes. Improvement of the teaching and assessment allotted time, poor teaching style and teaching materials, poor pre-course disseminated teaching resources, and the non-systematic delivery of the courses were suggested. This suggests that 90% of the students are satisfied or strongly agree with the effectiveness of the course materials and instruction.

### **Faculty perceptions of the SET**

Faculty perception is that faculty tend to “*somewhat agree*” that the SET outcomes reflect their teaching thus suggesting that the SET and FSET outcomes are somewhat correlated. The probability  $P(>4) = 64$  shows that 82% of the faculty somewhat agree that the outcomes of the SET accurately reflect their teaching.

The qualitative comment section provides faculty unrestricted opportunities to express their opinions about the implemented SET in the HEI. By using qualitative classification method (Bailey, 1994) after coding, the comments are classified into collections of meanings. Faculty recommendations for improving the case HEI's SET include the separation of the course materials and instructor surveys to reduce effect of one on the other, repeat questions to test bias, use more benchmarked questions and the SET questions should include formative and summative assessments. By suggesting that the course materials and instruction surveys should be separated due to the impact of one on the other, the faculty agree with Spooen & Mortelmans (2006) that a halo-effect exists between these ratings in the same SET document. This further emphasizes the high correlation found between the course materials and instruction correlation. Mean rating of  $3.45 \pm 1.16$  in question (2) suggests that the faculty neither agree nor disagree with the assertion that the questions in the SET are sufficient for evaluating the teaching skill of the faculty.

### **Faculty perceptions of alternative tools**

The questions in table 8 elicit faculty perceptions of implementing alternatives to SET survey. Question (3) explores faculty perception on the use of standardized teaching framework (Glaister, & Glaister, 2013) for evaluating faculty teaching quality. With a mean of  $3.35 \pm 1.16$ , 45% of the faculty strongly or somewhat agree with the use of a standardized framework for evaluating teaching over the SET. With a mean of  $1.91 \pm 0.90$ , the faculty disagree or are neutral to the use of independent consultants in the evaluation of effectiveness of teaching over the SET. A third of the faculty strongly disagree with the use of FSET over SET. The mean value of  $3.09 \pm 1.16$ . However, 45% of faculty agree or strongly agree with the use of the FSET over the SET. With a mean of  $2.27 \pm 1.05$ , faculty peer evaluations of teaching are not favoured by the faculty either. This suggests that the faculty prefer the use of SETs over other evaluative tools.

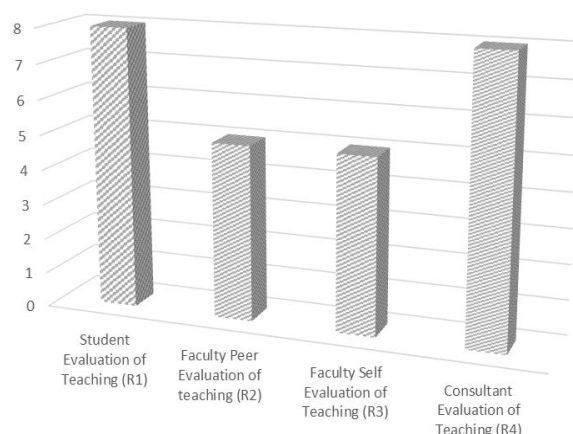
### **Faculty preferred evaluative tools**

Faculty also ranked SET, FSET, independent consultant evaluation, and peer evaluations in order of preference for PE. Figure 1 suggest that SET and use of independent consultants for evaluating teaching are equally preferred by faculty. FSET and peer evaluations are equally ranked below the SET and use of independent consultants. Faculty



justifications for the ranking includes students' experience of the teaching over time. Thus, faculty perceive that the graduate students' understanding of a "good lesson" is gained by students over their long-term experiences of teaching in higher education. Faculty also see the cost implication of using independent consultants with only a snapshot of teaching in the HEI as burdensome. Lack of experience of the local context by the independent consultants is also seen as a source of bias.

The HEI has one college with similar academic fields hence a singular teaching approach is adopted. Faculty suggest that the approach to evaluation needs to be benchmarked to specific standard for implementing FSET, independent reviews and peer evaluations. SET does not require a specific teaching framework hence in an absence of a standard, it coopts student perception of a good lesson. This finding runs contrary to the belief that seeing students as customers damage the quality of teaching, student learning and academic standards (Molesworth et al., 2009). Adoption of the SET with any/all the FSET, peer evaluations, or consultant evaluations would improve the robustness of the evaluation of teaching. Using only one of these tools and not the others may under-evaluate the effectiveness of the teaching.



**FIGURE 1**  
**ORDER OF PREFERENCE OF SET, FSET, CONSULTANT AND PEER EVALUATIONS**

### Factors challenging SET in PE

Table shows faculty perceived factors challenging the use of SET for faculty performance evaluation in the case HEI. In response to question 7, 64% of the faculty somewhat or strongly disagree that students are not sufficiently skilled to evaluate the teaching skills of the faculty. This conflicts with the findings of Mohammad (2001) and Gaillard et al. (2006) who viewed students' lack of training on course materials as an invalidating factor of SET outcomes. This approach suggests that the students must have curriculum and implementation skills to evaluate the effectiveness of the lessons. It is arguable that as customers and graduate students, they may be best positioned to know if the lessons are meeting their needs. 54% of faculty somewhat or strongly agree that the SET has a legitimate place as a consumer control tool and not as a PE tool. The response to question (9) suggests that 63% of faculty believe SET outcomes are more skewed by faculty charisma than the subject of the lesson. This aligns with the findings of Shah et al. (2020). Faculty charisma may go a long way in maintaining student attention but does not effectively address knowledge gained or the quality of the course materials.

81% of the faculty are either not sure, somewhat or strongly disagree that the lecture content has little effect on the student ratings. Combined, it indicates that subject content is important however, faculty charisma has a larger effect. The response to question 11 suggests that faculty are evenly split on their favourability of the use of the SET as an evaluative tool. The mean of 3.91 suggests student ratings of question (12) strongly indicate that the faculty believe that there are other important aspects of teaching skills which could not be evaluated by Likert scale ratings. This agrees with Wright (2005) who found that the SET may be inadequate for evaluation of teaching. The qualitative survey section allows students to make comments of desired improvements not captured in the ratings.

Faculty suggested that bias is positively introduced by friendly peers or negatively introduced due to faculty feud or disagreements. This suggests that peer reviews are skewed by internal conflicts and other biases. The perception is that peer evaluations could complement other tools used to inform PE and should not be used as a stand-alone evaluative tool. If objectively implemented, peer reviews are also believed to inform decisions on professional development.

### Disadvantage of the implemented SET

Table captures faculty perceptions of the disadvantages of implementing the SET. Response to questions (13), (14) and (15) suggest that faculty members of the HEI disagree or somewhat disagree that the implemented SET negatively affects their decisions in selecting teaching approach, relating with the students and the design of the course materials. Question (16) suggests that 18% of faculty somewhat or strongly agreeing that anonymous student response encourages silly and amusing response to the SET questions.

82% of the faculty members however somewhat or strongly agree that extreme responses skew the SET ratings (Dolnicar & Grun, 2009). Ratings of question (18) indicate that 45% of faculty believe that summative assessment negatively impact the outcomes of the SET however, 55% are either unsure or somewhat disagree that it has an impact. With a mean rating of 3.73, at least 73% of the faculty somewhat or strongly agree that the students' written in comments are generally constructive. 82% somewhat or strongly agree that the comment section is the most utilitarian section of the SET. 62% view it as most useful for developing faculty.

### Are FSET and SET outcomes: do they agree?

A simple regression analysis was used to determine the correlation between the outcomes of aspects of instruction in the SET and FSET for the 11 faculty. Table 4 shows the outcomes of the regression analysis. With  $r = 0.489$  and  $p = 0.403$ , SET and FSET outcomes are moderately positively correlated however, this relationship is not significant at  $\alpha = 0.05$  hence the model is not a good model and SET scores do not reflect the FSET ratings. This suggests that although faculty are satisfied with the SET outcomes, their ratings of the FSET only moderately reflect students' ratings of teaching Table 4.

Table 4 REGRESSION OF THE FACULTY AND STUDENT RATINGS OF THE COURSE SATISFACTION QUESTIONS							
	Unstandardized Coefficients		Standardized Coefficients	t	p	95.0% Confidence Interval for B	
	B	Std. Error	Beta			Lower Bound	Upper Bound
(Constant)	-5.032	9.839		-0.511	0.644	-36.343	26.279
SET	2.105	2.166	0.489	0.972	0.403	-4.788	8.999

Note: $r = 0.489$ , $\alpha = 0.05$ , Dependent Variable: FSET
--

## Discussion and Conclusion

### The SET and its outcomes

Outcomes of quantitative ratings of the SET measures of teaching effectiveness suggests that clear spikes indicating that faculty teaching are effective and ineffective where the teaching is unsatisfactory. The SET outcomes also show that the ratings are spread out, clearly delineating effective teaching and weak teaching. The findings are similar for the ratings of the course materials. The findings agree with Debroy et al., (2019) that students ranking of faculty is  $>4$ . The mean scores of students' course materials and instruction ratings agree with Spooren & Mortelmans (2016) that there is likely a halo-effect between the course and instruction ratings because they are both completed by the students in the same survey event. Outcomes of student ratings of the course materials and instructions taken in the same SET events suggest that the ratings of course materials and instruction are very strongly correlated. This suggests that effective instructors use effective teaching materials or students' perception of course materials and instruction are not delineated but multiplexed.

The findings provide stronger evidence supporting the findings of Spooren (2010) that course ratings and instruction ratings are similar. A note of caution is that the SET is implemented to assess students' opinions within the "*happiest period*" of the course where student opinions are untainted by the summative assessments. If the SET is implemented post-release of the summative assessment grades, student ratings of the SET are likely to be impacted by these assessments. By completing the comments box, only a small number of students volunteered opinions on how to improve teaching and course materials.

### Faculty perceptions of the SET

The findings suggest that faculty somewhat agree with the design of the SET and the outcomes for evaluating their teaching. However, faculty in agreement with Wright (2005) found that the SET insufficiently evaluate aspects of teaching. Thus, the faculty perceptions suggest that students' ratings in the SET accurately reflect their teaching (Beran and Rokosh, 2009; Idaka et al., 2006; Inko-Tariah, 2013). This finding negates the dissatisfaction expressed by Iyamu & Aduwa-Oglebaen (2005) and the challenges expressed by Iyam & Aduwa-Oglebaen, 2005; Mwachingambi & Wadesango, 2011; Boring et al., 2016; Hornstein, 2020 and Spooren et al, 2013.

Halo-effect seen in the current SET format and observed by Spooren & Mortelmans (2006) could be removed by separating the course materials and instruction surveys in the SET into independent surveys. Biases in the questions are also better reduced by ensuring that the SET questions support teaching improvements. Benchmarking of the questions and inclusion of summative and formative assessments questions should be embraced to improve the robustness of the SET. Use of standardized framework for evaluating teaching was somewhat supported over the simplistic and subjective outcomes of the SET. Faculty support for the use of FSET, independent consultant, and peer evaluations is indicated as hinging on the adoption of a teaching framework. Nonetheless, the faculty ranked the SET and FSET jointly higher than independent consultant, and peer evaluation. Faculty also suggested that peer evaluations are negatively impacted by internal conflicts, peer embellishments and other challenges. Financial burdens and evaluator insufficient

knowledge of the local context were indicated as challenges to adoption of independent consultants.

The findings indicate that the SET is not challenged by student's knowledge of a good lesson. The graduate students of the HEI understand what a "good lesson" is from inductive and deductive knowledge gained by students over their long-term experiences of teaching in higher education. This runs contrary to findings of Mohammad (2001) and Gaillard et al. (2006) viewed students' lack of training on course materials as an invalidating factor. It is arguable that as customers, graduate students may be best positioned to know the degree of effectiveness.

### **Are the FSET and SET outcomes related?**

Like the findings of Marsh (1987) and Beran & Violato, (2005), FSET and SET outcomes are moderately correlated suggesting that the graduate students in this HEI have some understanding of a good lesson and is moderately aligned with the faculty views. However, the robustness could be improved by broadening the questions designed to evaluate teaching effectiveness beyond the five questions used in the case HEI's SET.

### **Weaknesses**

The SET's 5 questions would not be sufficient to research teacher effectiveness. An SET template with questions testing teaching effectiveness would provide a broader understanding of the subject. The faculty population is 20 hence is insufficient for conducting a definite statistical study but an indicative one. Also, the number of female faculty is 4, thus limiting research related to gender biases. Agglomerating the data on the 20 faculty members to create an institution-wide perspective suggests that important information on the individual instructor may be lost. Micro-level studies of individual faculty effectiveness would be better served using interviews.

## **REFERENCE**

- Bailey, K. D. (1994). *Typologies and taxonomies: an introduction to classification techniques*. Sage.
- Corno, L. (2001). Self-regulated learning: A volitional analysis. In B. J. Zimmerman & D. H. Schunk (Eds.), *Self-regulated learning and academic achievement: Theory, research, and practice* (2nd ed., pp. 111–142). Mahwah, NJ: Erlbaum.
- Ginns, P., Prosser, M., & Barrie, S. (2007). Students' perceptions of teaching quality in higher education: the perspective of currently enrolled students. *Studies in Higher Education*, 32(5), 603–615.
- Hou, A. Y. C., Hill, C., Lin, A. F. Y., & Justiniano, D. (2022). The battle for legitimacy of 'student involvement' in external quality assurance of higher education from Asian QA perspectives—does policy contextualization matter? *Journal of Asian Public Policy*, 1-19.
- Johnson, R. (2000). The authority of the student evaluation questionnaire. *Teaching in Higher Education*, 5, 419-434.
- Keeley, R. G. (2015). Measurements of student and teacher perceptions of co-teaching models. *The Journal of Special Education Apprenticeship*, 4(1), 4.
- Lattuca, L. R., & Stark, J. S. (2009). *Shaping the college curriculum: Academic plans in context*. John Wiley & Sons.
- Marsh, H. W. (1984). Students' evaluations of university teaching: Dimensionality, reliability, validity, potential biases, and utility. *Journal of Educational Psychology*, 76, 707-757.
- Park, K. A., Bavis, J. P., & Nu, A. G. (2020). *Branching Paths: A Novel Teacher Evaluation Model for Faculty Development*.
- Theall, M. and Keller, J.M. (2017), Summary and Recommendations. *Teaching and Learning*, 2017: 109-112.
- Uwe Flick. (2009). *An introduction to qualitative research*. Sage.
- Vogt, W. Paul (1999). *Dictionary of statistics and methodology*. Sage: Thousand Oaks, California.

Indexed at, Google Scholar, Cross Ref

Voss, R., & Gruber, T. (2006). The desired teaching qualities of lecturers in higher education: A means end analysis. *Quality Assurance in Education*, 14(3), 217-242.

Wachtel, H. K. (1998). Student evaluation of college teaching effectiveness: A brief review. *Assessment & Evaluation in Higher Education*, 23, 191–212.

Zimmerman, B. J., (1998). Developing self-fulfilling cycles of academic regulation: An analysis of exemplary instructional models. In D. H. Schunk & B. J. Zimmerman (Eds.), *Developing self-regulated learners: From teaching to self-reflective practice* (pp. 1–19). New York: Guilford Press.

<b>Received:</b> 02-Oct-2024, Manuscript No. ASMJ-24-15305; <b>Editor assigned:</b> 07-Oct-2024, PreQC No. ASMJ-24-15305(PQ); <b>Reviewed:</b> 16- Oct-2024, QC No. ASMJ-24-15305; <b>Revised:</b> 25-Oct-2024, Manuscript No. ASMJ-24-15305(R); <b>Published:</b> 10-Nov-2024
--