

## **STUDENT EVALUATION OF TEACHING EFFECTIVENESS: EFL TEACHERS' AND STUDENTS' PERCEPTIONS**

**Tran Ngoc Bao Chau, MA**

*Can Tho University, Can Tho, Vietnam*

[fourleaf.clover1705@gmail.com](mailto:fourleaf.clover1705@gmail.com)

**Assoc. Prof. Truong Vien, PhD**

*Hue University of Foreign Languages, Hue, Vietnam*

[truongviensp@gmail.com](mailto:truongviensp@gmail.com)

---

**ABSTRACT:** *Student Evaluation of Teaching Effectiveness (SETE) is commonly used as a tool of evaluating teaching quality at educational institutions. In Vietnam, schools and universities have still imposed a restriction on using SETE and the issue on perceptions of SETE has yet to be resolved. Acknowledging this literature gap, the current study aims to investigate how EFL teachers and EFL students perceive SETE as well as to seek for possible similarities and differences between the two groups' perceptions with the participation of 130 teachers and 192 adult learners at an EFL language center in Southern Vietnam. Quantitative data were collected via questionnaires with two categories: (1) questionnaire for teachers, and (2) questionnaire for students. The findings revealed that Vietnamese EFL teachers and students were highly aware of SETE in relation to its necessity as well as potential benefits and the participants almost keep neutral viewpoint on validity and reliability of SETE. Also, the two groups acknowledged their frequent participation in SETE and expressed their preferences for certain SETE procedures. What's more, there were several similarities and differences between teachers' and students' perceptions. Eventually, the study puts forward some implications to make SETE activities more effective in the future. Further studies can focus on perceptions and practices of SETE in a different context or a comparison between SETE and another form of assessment.*

**KEY WORDS:** student evaluation of teaching effectiveness, SETE, teaching quality, course improvement, student feedback, forms of assessment

---

## **INTRODUCTION**

Nowadays, we are living in a world in which the job market is becoming more and more dynamic so it is vital that people make adaptations to better their work. Learning to learn

is considered a requisite for entry into most professions (Bridgestock, 2014). Also, Smith and Martin (2014) emphasize that reflective skills together with lifelong learning play a key role in making people professional at work. Thus, teachers are required to be reflective workers who make frequent reflection about their own teaching and improve teaching quality based on students' feedback since there will not be a development of career without the capability of reflection. Gaining evaluations from others such as colleagues, alumni or students is one of the ways assisting teachers to promote their reflective skills.

Student evaluation of teaching effectiveness (SETE or SET) has been one of the most popular tools used to enhance the accountability of various universities all over the world (De Witte & Rogge, 2011). Wachtel (1998), who published a brief review on student evaluation of college teaching effectiveness, mentions Remmers and Brandenburg as the pioneers of the preliminary research on SETE. According to Chen and Hoshower (2003), teaching effectiveness is even defined and measured in order to make a contribution to many decisions in higher education. What is more, that teachers' instruction is a decisive factor behind students' performance enables educators to make great attempts to enhance teaching assurance. As a consequence, the concern of improving teaching quality has proliferated so far (Ball & Forzani, 2009). A great number of institutions of higher education all over the world have been using SETE as a means of measuring teaching effectiveness (Zabaleta, 2007). Macfadyen et al. (2016) are in agreement with this viewpoint featuring an increase in the prevalence of SETE across higher education in an effort to guarantee course improvement and quality. In some cases, SETE is regarded as the only indicator of teaching quality so the schools where SETE is implemented have made efforts to develop optimum procedures and instruments so as to collect, analyze and interpret the data. Penny (2003) supposes that the convenience of this process has helped SETE gain its popularity over recent years. More importantly, according to Seldin (1993), it is students, significant stakeholders, who straightforwardly participate in the lessons and give insight into the instruction because "the opinions of those who eat the dinner should be considered if we want to know how it tastes" (p.40).

**Research questions:** The current research aims to answer the following questions:

1. How do the teachers perceive student evaluation of teaching effectiveness?
2. How do the students perceive student evaluation of teaching effectiveness?
3. What are the similarities and differences between teachers' and students' perceptions of student evaluation of teaching effectiveness?

## **LITERATURE**

Prior literature on main purposes of conducting SETE was introduced by some scholars (Chen & Hoshower, 2003; H W Marsh & Dunkin, 1992). Specifically, educators conduct SETE to reach four major targets: (1) to enhance teaching and course quality; (2) to make contribution to administrative decisions; (3) to help students select course units and lecturers; (4) to provide plenty of data for academic community in which SETE has still been under research. From another perspective, Kember, Leugn and Kwan (2002) introduce three major purposes of SETE, namely (1) to improve teaching quality, (2) to consolidate

administrative decisions (e.g. tenure or promotion), and (3) to demonstrate the responsibility of the institution. Calderon et al. (1996) claim that merely evaluating teaching effectiveness as well as relevant aspects of lecture is not a central objective of SETE. Instead, it is students' involvement, commitment and interest in their disciplines that should be fully appreciated.

In terms of potential benefits, SETE is confirmed to be advantageous to administrators (Cook-Sather, 2006), teachers (Smith, 2008) and students (Spencer & Schmelkin, 2002). Managers can collect student feedback for future reference such as making decisions of tenure or promotion. Moreover, student feedback on teaching effectiveness enables teachers to adapt their instruction and enhance their own growth and reflection. In addition, SETE proves beneficial to students by turning them into professional teacher watchers within the capability of making reasonable and sensible judgments about teaching. However, possible limitation on validity and reliability of SETE is of great concern as well (Yunker & Yunker, 2003). Many questions arise as to whether or not student evaluation is legitimate with respect to performance management and quality assurance (Stowell et al., 2012). A dearth of belief in reliability of SETE is attributed to some different elements such as perceived biases of kinds of student (Centra & Gaubatz, 2000), a reduction in response rates (Adams & Umbach, 2012) and other influential factors. It was almost 30 years ago when SETE was stated to be probably "the most thoroughly studied of all forms of personal evaluation" (Marsh, 1987, p. 369). It is noticed that the majority of studies on SETE focus on characteristics or validity and reliability of the employed questionnaire (Centra, 2003; Marsh, 2007); factors influencing student ratings and adequate evaluation of teaching quality (Brockx et al., 2011; Isely & Singh, 2005; Weinberg et al., 2009); but rarely do they give insights into perceptions of both teachers and students.

With regard to the practice of SETE, questionnaires have gained a good reputation as the best form of student evaluation (Blair & Noel, 2014) but there are various effective means of collecting students' opinion including one to one student interviews, e-mail, bulletin boards, students' diaries and informal comments. In general, it is advisable to promote an integration of different mechanisms with the intention of thoroughly aggregating student feedback (Brennan & Williams, 2004). The procedures of evaluation have evolved and been applied in 11 western colleges, according to McGee (1995), comprising four major forms of assessment: student, peer, self and administrative evaluation. There is an inclination to compare student evaluation with other forms of assessment like self-evaluation, peer evaluation, and alumni evaluation on the grounds that the standardized criterion of SETE is unsatisfactory.

Nevertheless, SETE is at times defined as a measure of students' expectations so it is supposed to address individual viewpoints rather than pedagogical knowledge. This is a reason why many people have still be skeptical about the validity and reliability of SETE (Ory, 2001). Merritt (2011) introduces some components affecting student evaluation consisting of teacher's smile, gesture, and other mannerism. As a consequence, it is possible that students underestimate teacher's knowledge, clarity, organization and other elements connected with good teaching performance on the assumption that the instructor

unintentionally drops a negative image into students' mind. What is more, there is a wide variety of factors leading to invalidity and unreliability of SETE entailing student's age (Sproule, 2002); student's gender (Heck et al., 2002); and degree of motivation and ability of students (Langbein, 2008). As an important element that institutes the reliability of SETE, the questionnaire designed for SETE is even entitled "happy form" (Penny, 2003, p. 400) and used as a measure of student's satisfaction. Thus, designing a questionnaire requires the architects to possess a complete understanding of what makes good and effective teaching (Knapper, 2001). In certain circumstances, if the instrument to collect SETE data is not valid and reliable enough, the main purpose of this evaluating activity will be converted into gaining high SETE scores instead of improving teaching quality (Simpson & Siguaw, 2000). These days, student evaluation of teaching effectiveness (SETE) is becoming more and more familiar with most researchers, educators and learners throughout the world. It was ascertained that the universal application of SETE stemmed from US, UK, Canada, Australia, and many European countries. That many Australian universities produced policies on the practice of online SETE was revealed by McCormack (2005). In Australia, the results of SETE were gathered from websites and used for faculty purposes like promotion or probation.

In the Vietnamese context, schools and universities have imposed a restriction on using student's ratings to evaluate teaching performance. As a matter of fact, despite its long-standing popularity in professional development and teaching evaluation in the world, SETE has just been common in Vietnam over the past 10 years. It is an explanatory incident due to the profound influence of Confucian Heritage Culture on the implementation of this procedure. In reality, Vietnamese students are supposed not to comment on their teachers' teaching so student evaluation is somehow alien and unpleasant for both evaluation givers and receivers. In Vietnam, it is enacted by the Ministry of Education and Training that the activities pertaining to SETE had better be looked up as useful sources to help teachers better their teaching and to determine teaching quality, especially at universities.

All things considered, it appears that SETE is not well-known to researchers and educators in Vietnamese context, particularly in Southern Vietnam. What is more, there are few studies focusing on SETE at a foreign language center. This current study is expected to address these gaps in the research of SETE aiming to investigate how EFL teachers and EFL students at a foreign language center perceive SETE. The results of this descriptive study could help enrich the source of studies on SETE and be valuable references for directors at foreign language centers to consider implementing SETE.

## **METHODOLOGY**

### **Participants**

The sample consisted of 130 teachers (40.4%) and 192 adult students (59.6%) from a foreign language center in the Mekong Delta, Vietnam. The first group were composed of 86 female teachers and 44 male teachers; and there are 108 girls and 84 boys in the second group. The method for selecting the participants in this study was convenience sampling. For a purpose of gathering the best information, the sample was chosen from the EFL classes in which SETE had frequently been applied. The teachers with experience in

conducting SETE were invited to join the study and they introduced some of their students who enthusiastically participated in SETE in previous courses. A short introduction of the research had been recommended to all of the selected teachers and students before they researched a consensus to give their answers to online questionnaires.

## **INSTRUMENT**

The questionnaire comprised 24 items aiming to investigate teachers' perception and students' perception towards SETE. It was adapted from those invented by Nasser and Fresko (2002) and Spencer and Schmelkin (2002) to suit the current study context. The first section included the participants' background information. The second section contained 24 closed-ended questions designed based on the format of a five-level Likert item ranging from (Strong disagree) to 5 (Strongly agree). The 24 questions were categorized into five clusters: the necessity of SETE (N); the practice of SETE (P); the benefits of SETE (B); the procedure of SETE (PRO); the validity and reliability of SETE (VR). The third section is relevant to possible difficulties and personal recommendations for a better situation. Table 1 below provides a summary of the questionnaires.

**Table 1.** Summary of the questionnaires

<b>Clusters</b>	<b>Number of items</b>
1. The necessity of SETE	4 items ( <i>Items 1-4</i> )
2. The practice of SETE	5 items ( <i>Items 5-9</i> )
3. The benefits of SETE	5 items ( <i>Items 10-14</i> )
4. The procedure of SETE	5 items ( <i>Items 15-19</i> )
5. The validity and reliability of SETE	5 items ( <i>Items 20-24</i> )

## **Procedure**

The procedure of this study was composed of three main phases, (1) piloting the instrument, (2) collecting data, (3) analyzing and interpreting the data. At first, 20 students and 10 teachers who were not from the sampling of this present study were responsible for checking whether the questionnaires could be clearly comprehended. Also, interview questions were piloted by two experienced and knowledgeable teachers who were not from the research population. After that, all participants including 192 adult students and 130 teachers at the EFL Language Center were explained the purpose of this study and asked for their participation in questionnaire completion prior to the distribution of questionnaires via e-mail. The researchers proceeded to manage this activity and then had all questionnaire answers gathered in the following week. The rest of the procedure focused on analyzing and interpreting the quantitative data.

### Data analysis

In this study, quantitative data were gathered from the participants' answers to the questionnaires and then they were encoded and analyzed with SPSS Amos version 24. First, preliminary analyses, including scale reliability and confirmatory factor analysis (CFA), were performed to investigate the reliability of the questionnaires.

To answer the first two research questions of EFL teachers' and students' perceptions towards SETE, the Descriptive Statistics Test was employed to measure minimum (Min), maximum (Max), mean scores (M) and standard deviations (SD) of every single item in the questionnaires. In order to answer the third research question of similarities and differences between the two groups' perceptions, Independent Samples T-Test was computed. The degree of agreement was based on the formula introduced by Kreinovich et al. (2019) in their book - Beyond Traditional Probabilistic Methods in Economic: *Interval width = (Maximum – Minimum) / Group = (5-1)/5 = 0.8*. Consequently, the interpretations of mean scores in case of a 5-point Likert scale are demonstrated as follows:

1.00 – 1.80: *Strongly disagree*                      2.61 – 3.40: *No idea*                      4.21 – 5.00:  
*Strongly agree*  
1.81 – 2.60: *Disagree*                                      3.41 – 4.20: *Agree*

## RESULTS

### Initial analysis of the questionnaire on perceptions of SETE

#### Scale Reliability Tests

In order to check the reliability of the questionnaire, a Scale Test was administered. The reliability analysis indicated that the questionnaire examining teachers' and students' perceptions of SETE was highly reliable to collect data for the current study. This conclusion was confirmed in support of the results of Scale Test (*Cronbach's Alpha > .7*). As can be seen from Table 2, all of the five factors, namely the Necessity of SETE, the Practice of SETE, the Benefits of SETE, the Procedure of SETE, and the Validity and Reliability of SETE, gained high values of Cronbach's Alpha ( $\alpha = .851$ ,  $\alpha = .870$ ,  $\alpha = .897$ ,  $\alpha = .826$ ,  $\alpha = .845$  respectively). Furthermore, the corrected item-total correlations are all greater than 0.3. Therefore, these scales are adequate and included in the CFA.

**Table 2.** Scale reliability test

Observed Variables	Scale Mean If Item Deleted	Scale Variance If Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha If Item Deleted
<i>The Necessity of SETE (N): Cronbach's Alpha = .851</i>				
Q1-N	12.39	3.826	.697	.808
Q2-N	12.42	3.827	.686	.813
Q3-N	12.48	3.870	.659	.825
Q4-N	12.37	3.923	.726	.798
<i>The Practice of SETE (P): Cronbach's Alpha = .870</i>				
Q5-P	15.90	8.426	.741	.831
Q6-P	15.78	8.721	.767	.824
Q7-P	15.85	9.037	.688	.844
Q8-P	15.62	9.639	.696	.843
Q9-P	15.62	10.111	.595	.865
<i>The Benefits of SETE (B): Cronbach's Alpha = .897</i>				
Q10-B	16.52	6.711	.743	.876
Q11-B	16.55	6.610	.762	.871
Q12-B	16.64	6.488	.752	.873
Q13-B	16.63	6.664	.759	.872
Q14-B	16.58	6.450	.720	.881
<i>The Procedure of SETE (PRO): Cronbach's Alpha = .826</i>				
Q15-PRO	15.82	7.700	.605	.799
Q16-PRO	15.75	8.014	.605	.797
Q17-PRO	15.68	8.018	.609	.796
Q18-PRO	15.62	8.305	.632	.790
Q19-PRO	15.55	8.248	.676	.779
<i>The Validity and Reliability of SETE (VR): Cronbach's Alpha = .845</i>				
Q20-VR	14.33	10.702	.682	.804
Q21-VR	14.66	10.331	.588	.836
Q22-VR	14.30	10.777	.659	.811
Q23-VR	14.18	11.452	.673	.810
Q24-VR	14.36	10.762	.683	.804

### Confirmatory Factor Analysis

In the questionnaire, five main factors incorporating 24 observed variables are generated with the intention of figuring out how EFL teachers and EFL students at a foreign language center in the Mekong Delta perceive SETE. Table 3 shows that the 24 variables are classified into five categories, with the total variance explained reaching 66.439%. What is more, the factor analysis exposes  $KMO = .869$  with the significant value of .000. The five dimensions of SETE were well defined in the factor structures from .681 to .886. The goodness of fit indices indicated that the data fitted the model well,  $CMIN/df = 1.233$ ,  $RMSEA = .03$ ,  $CFI = .95$  and  $PCLOSE = 1$ . Thus, the use of CFA in this current study is considered appropriate and these factors, as a consequence, can be utilized for the forthcoming analysis.

**Table 3.** Exploratory factor analysis of independent variables.

Observed Variables	Factor				
	1	2	3	4	5
Q10-B	.886				
Q11-B	.862				
Q12-B	.855				
Q13-B	.817				
Q14-B	.806				
Q5-P		.853			
Q6-P		.841			
Q7-P		.837			
Q8-P		.830			
Q9-P		.681			
Q20-VR			.850		
Q21-VR			.825		
Q22-VR			.763		
Q23-VR			.760		
Q24-VR			.693		
Q15-PRO				.837	
Q16-PRO				.763	
Q17-PRO				.750	
Q18-PRO				.738	
Q19-PRO				.722	
Q1-N					.863
Q2-N					.855
Q3-N					.835
Q4-N					.767
Eigenvalue	7.145	2.945	2.312	1.904	1.639
% of Variance	29.772	12.269	9.634	7.932	6.831
KMO					.869
Bartlett's Test			Chi-Square df Sig.		4024.649 276 .000



**Teachers' and students' perceptions of SETE**

A descriptive statistics test was computed to investigate teachers' and students' perceptions of SETE pertaining to necessity, practice, benefits, procedure, and validity and reliability of SETE. The results of this test are in charge of giving answers to the two research questions concerning how teachers and students perceive SETE in the context of an EFL language center. Table 4 below presents the results of the descriptive statistics test.

**Table 4.** Teachers' and students' perceptions of SETE

Item	Teacher N=130		Student N=192	
	M	SD	M	SD
<i>The Necessity of SETE</i>				
	4.1	.62	4.1	.65
1/ SETE is an essential activity in EFL classrooms.	4.2	.75	4.1	.79
2/ SETE can help the managers gain an overview on teaching and learning quality in general.	4.1	.71	4.2	.74
3/ Teachers will improve their teaching if they implement SETE in classes.	4.1	.71	4.1	.83
4/ The frequent application of SETE helps EFL learners study better and more effectively.	4.0	.82	4.0	.77
<i>The Practice of SETE</i>				
	3.7	.80	4.0	.69
5/ I join SETE via completing the questionnaires at the end of the course.	3.6	1.02	4.0	.86
6/ Many of students' opinions from SETE have been noticed and applied by the teacher.	3.5	1.09	4.0	.80
7/ I participate in SETE in almost all courses I have attended.	3.5	1.11	3.9	.95
8/ The SETE questionnaire combines 5-point Likert Scale questionnaire with open-ended questions in student evaluation form.	4.0	.85	4.0	.81
9/ When students give low evaluations, teachers adjust to improve their teaching.	3.9	.87	4.1	.79
<i>The Benefits of SETE</i>				
	4.1	.63	4.1	.63
10/ SETE assists the directors in perceiving teaching and learning conditions as a whole.	4.0	.79	4.1	.75
11/ The use of SETE demonstrates the accountability and great concern of the managers towards teaching and learning processes.	3.9	.75	4.1	.70
12/ SETE helps improve the quality of curriculum and teaching materials.	4.1	.74	4.1	.74
13/ SETE helps teachers gradually improve their professional knowledge and skills.	4.2	.75	4.2	.71
14/ SETE builds up students' confidence and independence when they can contribute their voices to the course improvement.	4.1	.81	4.1	.80
<i>The Procedure of SETE</i>				
	3.8	.61	3.9	.74

	4		7	
15/ There should be a team that is responsible for managing SETE at the English language center.	3.9 8	.84	4.1 0	.79
16/ SETE via electronic questionnaire is a good method.	3.7 6	.81	4.1 4	.83
17/ Informal talk is an effective way to figure out learners' opinions on the course.	3.9 8	.82	3.8 9	.99
18/ I support SETE carried out via the completion of questionnaires in the classroom at the end of the course.	3.7 8	.86	3.9 1	.97
19/ SETE should be conducted during the course.	3.7 2	.94	3.8 3	1.04
<i>The Validity and Reliability of SETE</i>				
	3.4 1	.81	3.7 2	.78
20/ There are not many differences in the results of SETE in a course.	3.3 3	1.05	3.7 8	.93
21/ Students have enough knowledge to judge the quality of instruction.	3.2 5	1.05	3.3 3	1.28
22/ Most students treat course evaluations seriously.	3.3 7	1.03	3.8 0	.97
23/ Good instructors get high course evaluations.	3.4 8	1.03	3.7 8	1.01
24/ The questionnaire administered in my foreign language center is a reliable mean for the evaluation of teaching.	3.6 1	.90	3.8 9	.85

### **The necessity of SETE**

With regard to the necessity of SETE, the total mean score of teachers' perception ( $M=4.14$ ,  $SD= .62$ ) and that of students' perception ( $M= 4.14$ ,  $SD= .65$ ) were observed to be reasonably high. In other words, the results indicate 130 teachers and 192 students at the EFL Language Center agree that SETE plays an essential role in their classes.

Particularly, Table 3 shows that EFL teachers expressed their fullest agreement on item 1 ( $M= 4.22$ ,  $SD= .75$ ) and EFL students offered their most approval for item 2 ( $M= 4.22$ ,  $SD= .74$ ). That is to say, the first group strongly agreed that SETE plays a significant role in their classes and the other gave their strong consent to the idea that the managers could grasp general information of teaching and learning quality thanks to SETE. As a whole, the total mean score of this factor and the mean scores of its four included items were at a high level ( $M > 4.0$ ), which emphasized the sample' firm belief in the necessity of SETE in their settings.

### **The practice of SETE**

The mean score of teachers' perception of the practice of SETE ( $M=3.78$ ,  $SD= .80$ ) was pretty high, which implies that EFL teachers reached their agreement on the suggested practices of SETE compared to their realistic application of SETE. Meanwhile, the mean score of students' perception ( $M=4.04$ ,  $SD= .69$ ) was high, which indicates that EFL

students agreed on the five suggestions for SETE practice to a greater extent in comparison with the surveyed teachers.

As can be seen from Table 4, to the greatest degree, the teachers expressed their general agreement on item 8 ( $M= 4.08$ ,  $SD= .85$ ) stating that a 5-point Likert Scale questionnaire accompanied by open-ended questions is used for SETE. From another perspective, the students showed their most approval for item 9 ( $M= 4.12$ ,  $SD = .79$ ) acknowledging teachers would be willing to adapt their teaching in case of receiving low evaluations. Generally, all mean scores in this factor ( $3.41 < M < 4.20$ ) reveal that the participants were aware of their practical application of SETE and agreed on the practices proposed in the questionnaire.

### **The benefits of SETE**

The mean score of teachers' perception of the benefits of SETE was high ( $M=4.11$ ,  $SD= .63$ ). In other terms, EFL teachers came to the same conclusion that SETE was considerably beneficial to three subjects, namely teachers, students and the board of managers. With regard to students' perception of SETE's benefits, the mean score was high ( $M=4.17$ ,  $SD=.63$ ), which means EFL students also reached an agreement that SETE was advantageous to teachers, students and managers of the foreign language center.

According to Table 5, item 13 gained the highest mean score from two groups of participants ( $M= 4.21$ ,  $SD= .75$ ;  $M = 4.21$ ,  $SD= .71$ ). In other words, both teachers and students in this survey strongly agreed that EFL teachers could make gradual improvement in their professional knowledge and skills as a result of SETE. The rest of the items illustrating other potential benefits of SETE achieved the mean scores at a high level of agreement (*from*  $M= 3.99$  *to*  $M = 4.19$ ), which places emphasis on both teachers' and students' belief in beneficial roles of SETE.

### **The procedure of SETE**

The mean score of teachers' perception of SETE procedure was relatively high ( $M=3.84$ ,  $SD= .61$ ) and that of students' perception was pretty high too ( $M=3.97$ ,  $SD= .74$ ). On average, a majority of the participants showed their agreement about the suggested procedural rules in the questionnaire.

To be more specific, from Table 6, teachers' responses reveal that the mean score of item 15 ( $M=3.98$ ,  $SD= .84$ ) and that of item 17 ( $M= 3.98$ ,  $SD= .82$ ) were the highest. It was interpreted that EFL teachers suggested having a team manage the SETE process and putting more informal talks into practice. Meanwhile, students' responses show that item 16 ( $M= 4.14$ ,  $SD = .83$ ) attained the highest mean score, which indicates that EFL students preferred using an electronic questionnaire as an evaluating instrument. Other statements regarding SETE procedure all received the accepted high level of agreement from the participants ( $3.41 < M < 4.20$ ). Hence, the study sample was noted to be under the agreement that most suggested procedural rules in the questionnaire were appropriate.

#### 4.2.5 The validity and reliability of SETE

The mean score of teachers' perception of the validity and reliability of SETE was at a medium level ( $M=3.41$ ,  $SD= .81$ ). It indicates that EFL teachers, to a certain extent, are not inclined to put their complete trust in students' comments. Meanwhile, the mean score of students' perception is relatively high ( $M= 3.72$ ,  $SD= .78$ ), which proves that the second group of participants in this study expressed more positive attitudes towards SETE's validity and reliability than the other.

To be more specific, Table 7 shows that the teachers expressed an agreement on item 24 to the greatest degree ( $M=3.61$ ,  $SD= .90$ ). In other words, they are under the assumption that the questionnaire used in their foreign language center is reliable enough for students to conduct SETE. In addition, they also agreed that good teachers tend to receive good feedback from their students ( $M= 3.48$ ,  $SD= 1.03$  - item 23). The mean scores of other items included in this factor, on the other hand, illustrate teachers' neutral opinions ( $M < 3.41$ ), which had an adverse effect on their belief in SETE's validity and reliability. Compared to the examined teachers, the students appeared to trust SETE's results more, which was verified by the high mean scores of most items (from  $M= 3.78$  to  $M= 3.89$ ) except that of item 21 ( $M= 3.33$ ,  $SD= 1.28$ ) stating that students are eligible for evaluating teaching quality. It means that the students remained neutral about this issue. All things considered, the results of this factor support the conclusion that teachers and students expressed their overall agreement when asked about the validity and reliability of SETE but the latter had more faith in valid and reliable evaluations than the former.

#### Similarities and differences between the two groups' perceptions

In order to answer the third research question, an Independent Sample T-Test was run. The data were analyzed via this test to investigate possible similarities and differences between the two groups' perceptions of SETE. The results of Levene's Test revealed that in terms of necessity, benefits and procedure of SETE, there was no significant difference between teachers' and students' perceptions ( $p= .969$  – *Necessity of SETE*;  $p= .362$  – *Benefits of SETE*;  $p= .085$  – *Procedure of SETE*). However, regarding practical application and validity and reliability of SETE, certain differences between the two groups' perceptions were observed ( $p= .002$  – *Practice*;  $p = .001$  – *Validity and reliability*).

Also, the findings of Descriptive Statistics Test indicated that EFL teachers in this study were the most aware of the necessity of SETE in their classes, which was proved by the highest mean score of their perception towards factor 1 ( $M=4.14$ ,  $SD= .62$ ). To the greatest extent, EFL students were in agreement about the benefits of SETE, which was deduced from the highest mean score of their perception towards factor 3 ( $M= 4.17$ ,  $SD= .63$ ). Another noticeable difference is that the teachers appreciated cluster 4- Procedure of SETE ( $M= 3.84$ ,  $SD= .61$ ), more than cluster 2- Practice of SETE ( $M= 3.78$ ,  $SD= .80$ ). On the other hand, the students overestimated cluster 2 ( $M=4.04$ ,  $SD= .69$ ) rather than cluster 4 ( $M=3.97$ ,  $SD= .74$ ). In terms of probable similarities, both groups of participants gained their consciousness of SETE's necessity to an almost similar degree of agreement ( $M=$

4.136 and  $M= 4.139$ ). Also, the two groups had a similar viewpoint as both teachers and students considered cluster 5- Validity and reliability of SETE- the least consented cluster ( $M= 3.41$  and  $M= 3.72$ ). Additionally, with the exception of factor 1 presenting the necessity of SETE, it should be noted that EFL students in this survey had a tendency to perceive the other factors at a higher level of agreement compared to EFL teachers. Moreover, all participants agreed on different aspects of SETE mentioned in the questionnaire, apart from the fact that the teachers seemed to keep neutral perspective on the validity and reliability of SETE ( $M= 3.41$ ,  $SD= .81$ ).

To sum up, the two groups of participants were inclined to have similar perceptions towards necessity, benefits and procedure of SETE but there seemed to be differences in their practice of SETE and their belief in valid and reliable student comments.

## **DISCUSSION**

The findings from the questionnaires showed that EFL teachers and EFL students believed in the essential role of SETE in their classes. This conclusion supported the hypothesis proposed by Zabaleta (2007) that considered SETE a crucial process for the institution. The highest mean scores for perception of the necessity indicated that both teachers and students were highly aware of the importance of SETE because of its apparent benefits to teachers, students and directors at the EFL Language Center in reality.

The results of the questionnaires also indicated that the participants frequently carried out SETE at the end of the course via completing a questionnaire. Moreover, teachers and students at the EFL Language Center acknowledged their experiences of using a 5-point Likert Scale questionnaire with open-ended questions. This finding confirmed what Blair and Valdez Noel (2014) found out, that is, a questionnaire using a rating scale has gained its popularity among various forms of student evaluation all over the world. One explanation for this is that questionnaires help teachers save a lot of time spent on SETE in their classes and provide students with specific criteria of evaluation. Another finding, based on the responses to the questionnaires, indicated that teachers were willing to modify their teaching methodologies when receiving low evaluations. This result was explained because no one except the teachers could directly read students' comments and then adjusted to improve their teaching. In this case, SETE can generate opportunities for teachers to reflect on their instruction, reform their own growth and predict learning needs as well. However, this statement was contrast to what Nasser and Fresko (2002) found in their study. The authors did research on teachers' perceptions of SETE and concluded that not many instructors at their college faculty took student ratings into account to adapt their instruction. Another finding regarding SETE practice showed that EFL teachers and EFL student were highly aware of SETE and put it into use on a regular basis, so there were not many considerable differences between perceptions and practices among the participants. It was because most teachers at the examined setting helped their students obtain appropriate understanding of SETE from the outset and always stimulated students' participation in the evaluating activity.

The finding was entirely consistent with a review of Marsh (1987), who claimed that SETE was beneficial to administrators, teachers and students. The participants believed that SETE proved responsibility and attention of the directors to teaching and learning quality. What is more, SETE was also noted to make students more confident to raise their voices in classes. In fact, student evaluations were sent to board of managers and the assessed teacher right after they had been assembled. Thus, students can feel that they virtually make a great contribution to teaching improvement and general development of the foreign language center in case they attend SETE.

When asked about their preferable procedures, the surveyed teachers appreciated informal talks as the most effective way of enabling students to express their opinions on the course. The finding was in agreement with a viewpoint stated by Spencer and Schmelkin (2002), who claimed that intimate conversations should be encouraged so that students would have great chance getting along well with their classmates and achieving gathered opinions. One explanation for this result is that an informal dialogue outside the classroom may create a secure and close-knit feeling that stimulates learners to naturally offer their comments. In contrast to teachers' most agreement on the application of informal talks, the participating students preferred the use of electronic questionnaires as a substitute for paper ones. It can be seen that EFL students, notably the youth, are active and creative enough to evaluate teachers' instruction via online questionnaires. Furthermore, the completion of electronic questionnaires can minimize students' insecure feelings and partially eliminate their fear of being prejudiced. The combination of different evaluating instruments (informal talks and electronic questionnaires) is opposed to the widespread use of paper questionnaires proposed by Blair and Valdez Noel (2014).

Finally, the findings from the questionnaires showed that almost all of the participants expressed their least approval for the validity and reliability of SETE. Particularly, both teachers and students moderately consented to valid and reliable results of SETE. In addition, the respondents in this survey agreed that a good teacher might gain high evaluations, which was inconsistent with prior studies (Feldman, 1977; Howard & Maxwell, 1980). The scholars suggested that if students are interested in the discipline or the instructor, their ratings are likely to be much higher. Another noticeable finding is that the participants kept neutral when asked about whether or not students have enough knowledge to judge their instructors. Meanwhile, Seldin (1993) firmly stated that students were important stakeholders in teaching improvement. It could be explained that students spend more time participating in class than an observing teacher but in some cases, different students can be biased witnesses who offer highly subjective viewpoints.

### **Implication**

In light of the major findings from this present study, five implications will be discussed as follows with the purpose of improving the quality of SETE and meeting the participants' demands.

Firstly, it is vital to help teachers and students raise awareness of the necessity of this evaluation activity. Meetings or conferences can be organized so that staff members and EFL teachers are well informed of the significance of SETE, its objectives, potential

benefits as well as thorough procedures. Above all, from the outset of the course, EFL students should be notified of in-depth information about SETE so as to grasp its process and make this activity much more effective.

Secondly, the board of managers and the leaders at the foreign language center should select appropriate forms of student evaluation and stimulate the procedure then. Hence, the assessment instrument should be carefully designed, which is expected to result in faith in the validity and reliability of the whole SETE process.

Thirdly, to collect SETE data, electronic questionnaires designed with specific evaluation criteria as well as informal talks should be considered. Therefore, a combination of intimate conversations, after-class conferences and an online questionnaire will be an ideal integration of measures instead of paper questionnaire delivery only.

Fourthly, it is recommended that the EFL language center should promote the practice of SETE via making it compulsory for all classes if possible. A reward-punishment system can be established so that students who enthusiastically participate in the SETE process will be complimented and teachers who are willing to join SETE will achieve pay rise or promotion. However, it is essential to make the participants aware of its significance at first and ready to join SETE without much pressure.

Finally, it is crucial to found a department at the foreign language center that takes responsibility for conducting SETE. The staff working for this department should be well trained to get on well with teachers and students in order to make the procedure run smoothly. Moreover, these friendly officers are supposed to be available all the time to contact the students as well as respond back to their comments if necessary.

## **CONCLUSION**

Main findings of this study revealed that the participants all expressed their agreement on different dimensions of SETE (necessity, practice, benefits, procedure, validity and reliability). Specifically, both EFL teachers and EFL students in this study were the most aware of the necessity of SETE in their classes. In addition, the perception of potential benefits was the second most agreeable to both sides. They acknowledged their frequent involvement in the SETE process commonly via questionnaire completion at the end of the course. Concerning favorable procedures of SETE, the first group preferred informal talks while the second supported the use of electronic questionnaires. Lastly, the participants in two separate groups agreed on the validity and reliability of SETE to a limited extent. These results proved that EFL teachers' and EFL students raised high awareness of SETE and took this evaluating activity into account. Hence, referring to the results of this current study, administrators at foreign language centers and headmasters at educational institutions can consider implementing SETE via different evaluation tools and bettering SETE with regards to procedural phases, validity and reliability.

## **Future Research**

Despite some positive findings of the current study, there are several limitations. First of all, the study was conducted with 130 teachers and 192 students, so the sample size ( $n=322$ )

was not large enough to generalize the results to all teachers and students at the foreign language center. Secondly, the outbreak of Coronavirus pandemic this year restrained the researchers from contacting the participants in person. As a consequence, the participants gave responses to online questionnaires and it took a quite long time to assemble enough data. Lastly, a few students ran into difficulty with finalizing the questionnaires. As a matter of fact, these students have recently experienced the SETE process and thus are still unfamiliar with this evaluation activity. Some participants revealed to spend much time reading and finishing the questionnaires.

After considering the limitations and the context of this study, some suggestions are withdrawn aiming to better the research on SETE in the future. To tackle the problems of online instruments, it is advisable to set up an official deadline for completing the questionnaires and the participants all deserve rewards for actively attending the survey. In order to provide stronger evidence on the findings, a longitudinal study involving all teachers and students at the foreign language center can be carried out. Furthermore, before inviting people to join the survey, the researchers should make sure that they all have knowledge or experience of SETE in order to save a lot of time spent on repeated explanation.

The results of this study can hopefully be used to enrich the future research into teachers' and students' perceptions of SETE. It is recommended that the next studies can be conducted to investigate both perceptions and practices of SETE in different contexts. Alternatively, to gain thorough assessment of teachers' instruction, further research is expected to combine student evaluation with other forms such as peer-evaluation and self-evaluation. Besides that, there is a need to explore perceptions of SETE in comparison with other forms of evaluation.

## References

- Adams, M. J. D., & Umbach, P. D. (2012). Nonresponse and Online Student Evaluations of Teaching: Understanding the Influence of Salience, Fatigue, and Academic Environments. *Research in Higher Education*. <https://doi.org/10.1007/s11162-011-9240-5>
- Ball, D. L., & Forzani, F. M. (2009). The work of teaching and the challenge for teacher education. *Journal of Teacher Education*. <https://doi.org/10.1177/0022487109348479>
- Blair, E., & Valdez Noel, K. (2014). Improving higher education practice through student evaluation systems: Is the student voice being heard? *Assessment and Evaluation in Higher Education*. <https://doi.org/10.1080/02602938.2013.875984>
- Brennan, J., & Williams, R. (2004). Collecting and using student feedback: a guide to good practice. *A Guide to Good Practice*.
- Brockx, B., Spooren, P., & Mortelmans, D. (2011). Taking the grading leniency story to the edge. The influence of student, teacher, and course characteristics on student evaluations of teaching in higher education. In *Educational Assessment, Evaluation and Accountability*. <https://doi.org/10.1007/s11092-011-9126-2>
- Calderon, T. G., Gabbin, A. L., & Green, B. P. (1996). Summary of promoting and



- evaluating effective teaching: American Accounting Association Teaching and Curriculum Section: Promoting and evaluating teaching effectiveness committee. *Journal of Accounting Education*. [https://doi.org/10.1016/0748-5751\(96\)00020-6](https://doi.org/10.1016/0748-5751(96)00020-6)
- Centra, J. A. (2003). Will teachers receive higher student evaluations by giving higher grades and less course work? *Research in Higher Education*. <https://doi.org/10.1023/A:1025492407752>
- Centra, J. A., & Gaubatz, N. B. (2000). Is there gender bias in student evaluations of teaching? *Journal of Higher Education*. <https://doi.org/10.1080/00221546.2000.11780814>
- Chen, Y., & Hoshower, L. B. (2003). Student evaluation of teaching effectiveness: An assessment of student perception and motivation. *Assessment and Evaluation in Higher Education*. <https://doi.org/10.1080/02602930301683>
- Cook-Sather, A. (2006). Sound, presence, and power: “Student Voice” in educational research and reform. *Curriculum Inquiry*. <https://doi.org/10.1111/j.1467-873X.2006.00363.x>
- De Witte, K., & Rogge, N. (2011). Accounting for exogenous influences in performance evaluations of teachers. *Economics of Education Review*. <https://doi.org/10.1016/j.econedurev.2011.02.002>
- Feldman, K. A. (1977). Consistency and variability among college students in rating their teachers and courses: A review and analysis. In *Research in Higher Education*. <https://doi.org/10.1007/BF00991288>
- Heck, J. L., Todd, J., & Finn, D. (2002). Is Student Performance Enhanced By Perceived Teaching Quality? *Journal of Financial Education*.
- Howard, G. S., & Maxwell, S. E. (1980). Correlation between student satisfaction and grades: A case of mistaken causation? *Journal of Educational Psychology*. <https://doi.org/10.1037/0022-0663.72.6.810>
- Isely, P., & Singh, H. (2005). Do higher grades lead to favorable student evaluations? *Journal of Economic Education*. <https://doi.org/10.3200/JECE.36.1.29-42>
- Kember, D., Leung, D., & Kwan, K. (2002). Does the use of student feedback questionnaires improve the overall quality of teaching? *Assessment and Evaluation in Higher Education*, 27, 411-425.
- Knapper, C. (2001). Broadening our approach to teaching evaluation. *New Directions for Teaching and Learning*. <https://doi.org/10.1002/tl.32>
- Kreinovich, V., Thach, N. N., Trung, N. D., & Van Thanh, D. (Eds.). (2018). *Beyond Traditional Probabilistic Methods in Economics* (Vol. 809). Springer.
- Langbein, L. (2008). Management by results: Student evaluation of faculty teaching and the mis-measurement of performance. *Economics of Education Review*. <https://doi.org/10.1016/j.econedurev.2006.12.003>
- Macfadyen, L. P., Dawson, S., Prest, S., & Gašević, D. (2016). Whose feedback? A multilevel analysis of student completion of end-of-term teaching evaluations. *Assessment and Evaluation in Higher Education*. <https://doi.org/10.1080/02602938.2015.1044421>
- Marsh, H W, & Dunkin, M. (1992). Students’ evaluations of university teaching: a multi-dimensional perspective. *Higher Education: Handbook of Theory and Research*.

---

*Volume VIII.*

- Marsh, Herbert W. (1987). Students' evaluations of University teaching: Research findings, methodological issues, and directions for future research. In *International Journal of Educational Research*. [https://doi.org/10.1016/0883-0355\(87\)90001-2](https://doi.org/10.1016/0883-0355(87)90001-2)
- Marsh, Herbert W. (2007). Students' Evaluations of University Teaching: Dimensionality, Reliability, Validity, Potential Biases and Usefulness. In *The Scholarship of Teaching and Learning in Higher Education: An Evidence-Based Perspective*. [https://doi.org/10.1007/1-4020-5742-3\\_9](https://doi.org/10.1007/1-4020-5742-3_9)
- McCormack, C. (2005). Reconceptualizing student evaluation of teaching: An ethical framework for changing times. *Assessment and Evaluation in Higher Education*. <https://doi.org/10.1080/02602930500186925>
- McGee, R. L. (1995). Faculty evaluation procedures in 11 western community colleges. *Community College Journal of Research and Practice*. <https://doi.org/10.1080/1066892950190406>
- Merritt, D. J. (2011). Bias, the Brain, and Student Evaluations of Teaching. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.963196>
- Nasser, F., & Fresko, B. (2002). Faculty views of student evaluation of college teaching. *Assessment and Evaluation in Higher Education*. <https://doi.org/10.1080/02602930220128751>
- Ory, J. C. (2001). Faculty thoughts and concerns about student ratings. *New Directions for Teaching and Learning*. <https://doi.org/10.1002/tl.23>
- Penny, A. R. (2003). Changing the Agenda for Research into Students' Views about University Teaching: Four shortcomings of SRT research. *Teaching in Higher Education*. <https://doi.org/10.1080/13562510309396>
- Seldin, P. (1993). The use and abuse of student ratings of professors. *The Chronicle of Higher Education*.
- Simpson, P. M., & Siguaw, J. A. (2000). Student Evaluations of Teaching: An Exploratory Study of the Faculty Response. *Journal of Marketing Education*. <https://doi.org/10.1177/0273475300223004>
- Smith, C. (2008). Building effectiveness in teaching through targeted evaluation and response: Connecting evaluation to teaching improvement in higher education. *Assessment and Evaluation in Higher Education*. <https://doi.org/10.1080/02602930701698942>
- Smith, S., & Martin, J. (2014). Practitioner capability Supporting critical reflection during workbased placement - A pilot study. *Higher Education, Skills and Work-Based Learning*. <https://doi.org/10.1108/HESWBL-04-2014-0009>
- Spencer, K. J., & Schmelkin, L. P. (2002). Student perspectives on teaching and its evaluation. *Assessment and Evaluation in Higher Education*. <https://doi.org/10.1080/0260293022000009285>
- Sproule, R. (2002). The underdetermination of instructor performance by data from the student evaluation of teaching. *Economics of Education Review*. [https://doi.org/10.1016/S0272-7757\(01\)00025-5](https://doi.org/10.1016/S0272-7757(01)00025-5)
- Stowell, J. R., Addison, W. E., & Smith, J. L. (2012). Comparison of online and classroom-based student evaluations of instruction. *Assessment and Evaluation in Higher*

*Education*. <https://doi.org/10.1080/02602938.2010.545869>

- Wachtel, H. K. (1998). Student evaluation of college teaching effectiveness: A brief review. In *Assessment and Evaluation in Higher Education*. <https://doi.org/10.1080/0260293980230207>
- Weinberg, B., Hashimoto, M., & Fleisher, B. (2009). Evaluating teaching in higher education. *Journal of Economic Education*. <https://doi.org/10.3200/JECE.40.3.227-261>
- Yunker, P. J., & Yunker, J. A. (2003). Are Student Evaluations of Teaching Valid? Evidence From an Analytical Business Core Course. *Journal of Education for Business*. <https://doi.org/10.1080/08832320309598619>
- Zabaleta, F. (2007). The use and misuse of student evaluations of teaching. *Teaching in Higher Education*. <https://doi.org/10.1080/13562510601102131>