Indian Journal of Information Sources and Services
ISSN: 2231-6094 (P) Vol.16, No.2, 2026, pp.752-766

DOI: https://doi.org/10.51983/ijiss-2026.16.2.76

Looking at Colleges and Universities Through the Eyes of Student Evaluations of Teaching Effectiveness (SETE)

Abdellateef Alqawasmi[1], Hamzeh Dodeen[2], Najeh Rajeh Alsalhi[3*], Sami Al-Qatawneh[4] and Ahmad Mohammad Alzoubi[5]

[1]Assistant Professor, College of Education, Humanities and Social Sciences, Al Ain University, Abu Dhabi, United Arab Emirates

[2]Professor, Department of Cognitive Sciences, Psychology Program, College of Humanities and Social Sciences, United Arab Emirates University (UAEU), Abu Dhabi, United Arab Emirates

[3*]Assistant Professor, Department of Education, College of Arts, Humanities and Social Sciences, University of Sharjah, Sharjah, United Arab Emirates

[4]Assistant Professor, Department of Education, College of Arts, Humanities, and Social Sciences, University of Sharjah, Sharjah, United Arab Emirates

[5]Associate Professor, College of Humanities and Sciences, Ajman University, Ajman, United Arab Emirates; Department of Psychology and Special Education, Al Balqa Applied University, Al-Salt, Balqa Governorate, Jordan

E-mail: [1]abdellateef.alqawasmi@aau.ac.ae, [2]hdodeen@uaeu.ac.ae, [3]nalsalhi@sharjah.ac.ae, [4]salqatawneh@sharjah.ac.ae, [5]a.alzubi@ajman.ac.ae

ORCID: [1]https://orcid.org/0000-0002-6127-131X, [2]https://orcid.org/0000-0002-9761-1988, [3]https://orcid.org/0000-0002-8807-4527, [4]https://orcid.org/0000-0003-0728-6761, [5]https://orcid.org/0000-0002-5550-130X

(Received 10 April 2026; Revised 15 May 2026, Accepted 22 May 2026; Available online 05 June 2026)

Abstract - **The evaluations by students on the effectiveness of their lecturers are key indicators in the process of faculty development and quality assurance in higher education. It is therefore important that the factors affecting the students' perception of teaching effectiveness be analyzed with an aim of enhancing teaching effectiveness. The current study analyzes the effect of the three student variables –academic year, GPA, and class size -on their evaluation of the teaching effectiveness using a descriptive research design. Data was collected from 713 university undergraduate students in different fields using an instrument derived from the student teaching assessment instrument (SETI). Findings indicated that first year students gave the highest rating for teaching effectiveness (Mean = 126.19, SD = 18.42), third year students the lowest (Mean = 118.48, SD = 20.69). Class size also had an effect such that those in small classes (1–15) students gave significantly higher rating (Mean = 128.67, SD = 19.91). GPA did not have significant relationship with the SETE scores ($F = 0.33$, $p > 0.05$). The reliability analysis revealed the consistency of the SETI measure with a high value of Cronbach's alpha at 0.97. The results highlight the need to take into account the maturity level of students and the environment when interpreting SETE outcomes. It is recommended that SETE must be understood based on its context, particularly from the cultural perspective, such as that of the Gulf region, where institutions must adjust their evaluation of teaching performance by accommodating demographic and environmental differences.**

Keywords: **Students, Evaluation, Teaching, Effectiveness (SETE), Higher Institution**

I. INTRODUCTION

The last decades of the 20th century and in the first decade on this one, we see an increasing use of academic performance evaluation in academies. The best example of this relationship with critical thinking has been the massive acceptance of quality standards in education, which require a complete and integrated audit of each ingredient within the teaching-learning process. There has indeed been a great deal of controversy concerning the interpretation of teacher grades. Only few studies focus specifically on grade inflation in relation to students' Self and Other ratings. Since this is an important task in academic life, it has been heavily worked on. However, that focusing on teachers' quality and originality is essential for better education in general. key features of successful schools. This is an indication that schools are truly desirous of improving learning environment. A crucial consideration is student evaluation of teacher effectiveness (SETE). SETE is becoming a typical way to judge the quality of teaching in higher education. Some ways to check how well higher education is educating are peer observations, supervisor appraisals, and reviews at the institutional level. Which is say that supervisor evaluations look at departmental needs and bigger teaching goals, while peer assessments focus on how well the teacher's methods and content delivery fit with the goals. Nonetheless, one of the most common and easy ways to judge how well a teacher is doing is to use student assessments (Ellett & Teddlie, 2003). Their findings illuminate aspects such as student-instructor interaction, instructional clarity,

fairness, and participation that may not be fully apparent through supervisor or peer evaluation. It is best to employ a multi-faceted approach to evaluation, but it is important to include student feedback so that you can get their point of view on how real and immediate the learning experiences.

Students take part in this process by providing thorough feedback on the class, the teacher's teaching style, and the depth of the course material. Because of this process, students contribute in shaping the educational terrain. By giving people, a chance to reflect on where we are in education, SETE will foster a growth mindset and a drive for excellence. To argues this is a necessary process for teachers, who are now improving their own art and in turn, the progress of their students. In the recent past, SETE has emerged as a widely used method to monitor and enhance the quality of teaching in universities. Surveys are increasingly used as an instrument for teachers' assessment. Semantic research shows that the SETE system can be a useful tool to improve teaching dynamics, methods and course design. There is a large literature on various characteristics and elements influencing SETE. These include factors related to the student such as gender, year of study, class size and performance in school. A few study have investigated the relationship between student sex and SETE, though findings were mixed. The investigation begins with a reflection on the difficulties of gender relations in education, and this picture shows its nuances in a multi-faceted way. This research explores the influence of student gender on SETE scores, focusing on biases and perceptions that may bias academic assessment. Grades have also been the subject of interest in educational research on the relationship of SETE to grades and whether and how students' academic performance may impact their teaching quality assessments. This complex relationship has findings indicating direct and indirect associations. The final factor that this study addresses is the class size. Again, it has been observed that SETE and class size were extensively researched in education (Mittal & Gera, 2015).

The study demonstrates that class size can influence students' evaluations of the teacher's performance. This effect can vary in appearance and magnitude based on factors such as the number of teaching assistants, the difficulty of the subject, the quality of instruction, the discipline, and the overall student-to-faculty ratio. More research in the relevant literature is needed to improve our understanding of how successful higher education instructors are. New study on teacher effectiveness and competence is critically needed, A more complete strategy to evaluating educators ought to be applied by all university college departments. It is vital to add contextual aspects when analyzing outcomes from SETE evaluations. From the perspective of the student, the value of this study becomes clear since it covers a fundamental part of education: the professional abilities required of university professors. Given that aiding faculty members in establishing their own professional knowledge is a primary objective of the school, this technique takes on considerable relevance. Academic year, gender, class size, and grade point average (GPA) were chosen as student-related characteristics because of the expanding quantity of empirical research demonstrating their impact on students' judgments of the efficacy of education (Uttl, 2024). Further than simple demographic statistics, these traits influence people's evaluations through interactions with motivational, contextual, and cognitive components (Beran et al., 2005). Research indicates that when students advance in their academic journeys and encounter many pedagogical approaches, with cultivate increasingly critical perspectives on the profession at large. Academically successful students may be more attuned to the clarity, organisation, and fairness of their lessons due to the relationship between GPA and evaluative sensitivity. The size of the class also has a big effect on how interested, involved, and responsive students think their teachers. While extensive research exists regarding these linkages in the global context, there is a notable deficiency of empirical studies examining Arab higher education environments and the influence of these factors on student evaluations (Garcia, 2025). To address this information gap, this study systematically investigates the influence of these student-related factors within a university context in the Gulf. This study investigates the influence of several significant and interrelated variables on SETE. Examining these links can help one better understand the SETE. Both the evaluation technique and the conclusions generated from it are enhanced by this, and the evaluation outputs are better validated. The impacts of gender, grade level, class size, and semester on SETE are studied in this research.

The UAE is a wonderful example because it wants to change and modernize its university system as part of its long-term plans. The UAE has placed an emphasis on educational quality, accreditation, and learner-centered pedagogies; as a result, students' opinions of instructors' efficacy may be evaluated within this framework, which is always growing. Compared to Western contexts, where SETE has been widely examined, the Gulf area has a distinct shortage of empirical investigation on the topic. To gain a more nuanced and culturally sensitive understanding of student assessments in Arab higher education, we aim to provide insights specific to this context by focusing on a university in the United Arab Emirates. By developing the Student Evaluation of Teaching Effectiveness (SETE) paradigm inside a Gulf university setting, this study fills a need in the literature that is mostly based in Western contexts. In contrast to previous theoretical models, this one takes a look at how gender, academic year, grade point average, and class size all have a role in teacher evaluations all at once. Academic policymakers and educators may benefit from the study's interpretation of SETE data (Centra, 2003). Better and more equitable assessments, as well as more robust programs to cultivate diversity-aware educators, are all possible outcomes of schools' increased awareness of the ways in which students' backgrounds impact their learning.

1.1 Issue with the Research

Student Assessment of Teaching Effectiveness (SETE) is of great interest to higher education policymakers due to the importance of student voices. However, research on the factors that influence university students' perceptions of their professors' classroom effectiveness is scarce. This research aims to examine how university students view their lecturers. In their study, emphasized the importance of investigating how diverse student and teacher attributes affect teaching effectiveness. Improving both student achievement and teacher quality requires increased teacher effectiveness, according to this study. Further research has also addressed the nature and function of teaching effectiveness in higher education institutions (Kocsis & Pusztai, 2025). According to this study, university students should be able to evaluate all departments, not just their own.

1.2 Objectives of the Study

University students' subjective ratings of instructors' pedagogical efficacy (SETE) are the principal focus of this research. The research tries to acquire a better knowledge of the elements that influence student input by looking at aspects like gender, academic standing, grade point average, and class size. Also, by underlining the need of multidimensional evaluation frameworks that take diverse student characteristics into consideration, the study intends to fill a vacuum in the literature. The ultimate purpose of these studies is to assist institutions enhance the quality of instruction across disciplines by building evaluation systems that are more equitable and responsive to instructors' demands.

1.3 Relevance of the Study

Students at the university desire a set of professional characteristics in their professors, and this study aims to determine the factors. Its focused-on scaffolding of education. This vital work becomes even more important because colleges can't function without investing in their teachers' professional development. Furthermore, this work analyses students' SETE scores in terms of their gender, year of study and class. Self-reflection, pedagogical adjusting, as well as student-oriented learning contribute to enhancing the teaching of their courses by applying the SETE method. In this article, In this describe the technique SETE, show some educational contributions and explain how its application could improve the quality of university teachers.

1.4 Study Questions

This research is guided by the following exploratory question:

What is the effect of demographic characteristics (gender, academic year level, GPA) and environmental characteristics (e.g., class size) on students' ratings of teaching effectiveness in higher education?

In order to probe this elementary question, the investigation raises the following more specific questions:

RQ1: Is there any difference between male and female students in their perceived of the quality Teaching effectiveness/quality ratings, and if so, would gender beliefs may play role on Expectations for these perceptions?

RQ2: What role does the students' academic year play in terms of their expectations and appraisals of teaching? Does progression through university lead to a more critical appraisal?

RQ3: How are students' GPA related to their perception of instructors after controlling for grade satisfaction?

RQ4: What is the role of the class size on the way students evaluate teaching effectiveness? Are there smaller groups that rate lower?

RQ5: Are there interactions effects—e.g., between sex and year or GPA and class size—that contribute to a more nuanced understanding of student rating patterns?

Its purpose is to identify issues of statistical inequities and provide perspective for universities on how strong their manipulation and judgement of teaching methodology and evaluation systems may lead to student contributions.

This research is organized into the following sections: Section 1 explains the introduction of the topic, and also defines the issue with the research, the objectives of the study, the relevance of the study, and the study questions. Section 2 explained about literature review consists of Research gap. Section 3 explained the theoretical framework, followed by gender, Academic year of the student, Grades, class size, Impact on perceived teaching quality, discipline-specific effects, level of study, threshold effects, longitudinal and contextual factors, and SEM Results. Section 4 explained methodology followed by architecture diagram, Research Design, population, Samples, ethical considerations, Instruments. Section 5 explained about results such as participants, design tools and Techniques, Reliability analysis, Gender differences, Academic year of student differences, grade differences. Section 6 explained the discussion, implications for practice and policy, limitations of the study and Section 7 explained the conclusion of this research.

II. LITERATURE REVIEW

SETE are here to support in making college better by making the teaching better. It enables the learners to reflect on what are learning and how like their lecturers to teach. Also, SETE will assist teachers in exploring how can make lessons even more effective to ensure that everyone has a good time learning. The paper also highlights evolution in research related to student evaluation of teaching effectiveness (SETE) in so far as little nuance has been given to the need for improvements to comprehensive system-wide

approaches, a need articulated by academics throughout this body of work, limitations have been acknowledged and more robust models developed that facilitate the use of student feedback to influence education practice at both faculty (individual) and institutional levels within higher education (Darling-Hammond et al., 2020). As noted by, amongst the reliable and consistent ways of assessing performance of university teachers and professional/technical competence depend on students' evaluations are one of them (Marsh, 2007). Evaluating college professors by their students is an important principle in the education system, particularly as it pertains to exams. The significance of SETE for the HAM The significance of SETE for the HEI culture can be illustrated by some qualitative aspects (Centra & Gaubatz, 2000).

The Student Evaluation of Teaching Effectiveness (SETE) paradigm bases this research with constructivism and behavioral views. SETE thinks students can appraise classroom education as learners. These exams evaluate instructional methods, class structure, comprehension, participation, and response. Gender, academic year, GPA, and class size are decided by theory and research. Teacher efficacy gender biases may be influenced by preconceptions. Students evaluate their own and others' instruction throughout the year depending on their growth and maturity. The study whether GPA, a popular academic accomplishment metric, affects students' judgments. Most teaching success models depend on interactions, feedback, and engagement, which class size can affect. Teaching efficacy in higher education depends on persons and context (Francisco et al., 2020). The student-centered evaluation approach in this study integrates these variables to better understand this.

The Teaching Effectiveness Through Students' Evaluation (SETE), which is an all-inclusive method with a special emphasis on the students' perspectives, has been gaining momentum. Several aspects about the student body have an impact on teaching effectiveness (Dervenis et al., 2026). Considerations include the gender, academic level, grade point average (GPA), and class size of students. In order to streamline the literature review for paper and make it concise, it is important to narrow down on synthesizing the relevant studies and removing the irrelevant references. The following approach can be used based on the context of the document: The importance of SETE in assessing the teaching effectiveness has become more significant than ever before. At first, gender, class size, and GPA had major implications in assessing the impact of gender biases and class sizes on teaching effectiveness. However, recent studies indicate a need to consider several other aspects (Gortney et al., 2026). Factors Affecting SETE: Previous studies have demonstrated the importance of class size on SETE, as small class sizes enhance the interactions between teachers and students, which increases the likelihood of positive evaluations. Academic year and GPA are also two important factors, since advanced students tend to be more critical and demanding. Although GPA was not one of the predictive factors of SETE in study, it can be justified by the notion that students become more mature and sophisticated with time.

Contextual Perspectives: A significant gap in SETE literature pertains to the Gulf and Arab regions because the literature is mostly based on Western contexts. Research, for instance, has shown that different cultures might affect student evaluations. In this regard, the current paper helps fill the identified gap by analyzing the UAE university context as well as studying the impact of gender, year of studies, grade point average, and class size on SETE. Necessity of a Context-Sensitive Approach: Another important lesson learned from the literature concerns the necessity of taking a context-sensitive approach to studying SETE as it should take into consideration particular features of the educational process and cultural peculiarities. In the present case, SETE is emphasized as a helpful tool that provides important information about teaching effectiveness; however, it should not be taken at face value without considering such factors as class size, gender, and progress of students through academic years. Summarizing Literature Thematically: In order to make the literature review more efficient and avoid repetition, it is necessary to summarize the findings thematically (Mahalli, 2026).

The most recent research efforts have been made in exploring various aspects of teaching quality, student evaluation, and other factors that affect the results obtained by students. One of them addresses issues regarding how primary school pupils comprehend cosmic phenomena, thus providing a new perspective into the evaluation of student involvement in complex disciplines and the importance of context in assessing educational efficiency. The work conducted by (Donado and Zerpa 2026) considers the interrelation between the quality of teaching and students' morality development stages, paying particular attention to the importance of using such concepts as TBC and DIT (Chytry et al., 2025).

These findings are consistent with the current study, which concentrates on the effects of variables such as the academic year and GPA on SETE. Moreover, looked at the issues related to teacher resilience among pre-service teachers in Myanmar, stressing the necessity of teacher well-being and its effect on instruction quality. It is connected with SETE studies that consider teachers' involvement and effectiveness in their work (Ergün et al., 2026; Phyo & Kopp, 2026). Within the domain of educational marketing, analyzed the effects of social media on student decision-making processes at a higher education level. This research is quite significant in terms of understanding external factors affecting students' perception and assessment of teaching quality. Finally, emphasized the importance of immersion in cultural education. Specifically, these scholars studied the role of university-school-community partnership in decolonizing teacher education and the changes it brings to teaching practices (Paliwal & Singh, 2026). These studies reveal the ongoing process of development within the field of teaching quality and SETE studies, thus, demonstrating that

multi-faceted approaches to evaluating teaching quality should be considered by researchers (Fauzi et al., 2025).

Research Gap

- Student engagement: SETE supports students in grounding themselves with achievable goals and actionable steps for success which is critical to academic success.
- Peers review of teaching: SETE is a component of review by peers of the process of teaching which stimulates synergistic thinking about teaching and learning.
- Setting the tone: SETE can set a vending point of an instructor's behavior and manners in class setting to decrease uncivility.
- Structure of the course: Structure of the course impacts prospective student performance, motivation and persistence. Conceptual understanding, synthesis of concepts and knowledge can be achieved for students using learning objectives, activities and assessments.
- SETE highlights the need to turn our lens back on student learning and utilize systematic inquiry to advance quality in higher education.

III. THEORETICAL FRAMEWORK

Student evaluation of teacher efficacy (SETE) underpins this study. This method is based on student-centered learning, where students assess their own progress, and constructivist learning theories. In some institutional settings, student ratings can reliably indicate how well a class would do. SETE frameworks combine clarity, structure, involvement, and fairness to attain educational goals. Students like these elements as class participation assesses instructional effectiveness.

The size of the class, the year of school, and the students' GPA all affect how feel about their teachers and class. The grade point average (GPA) may show how effectively you take criticism and learn, but older students may rate instructors differently since are more mature academically. SETE success depends on class size, which impacts engagement, interaction, and instructor reaction. These international studies are informative, but may not reflect the cultural, institutional and pedagogical circumstances of higher education in the Middle East. This research adds to a small regional literature on SETE in Arab universities that seeks to redress this gap. Yet few studies have specifically examined how features of the student (academic year, GPA, and class size) influence perceptions of teaching from this perspective. By investigating these dimensions in a Gulf university setting, this study seeks to add to the existing literature through providing more localized insights into how students perceive classroom teaching and learning effectiveness within their education landscape.

3.1 Gender

Some studies find that female faculty are evaluated less favorably in STEM than male faculty. These differences have been attributed to gender stereotypes and biases that influence students' views of teaching effectiveness. For example, in a study by focused on an online course, students who thought their instructor was male rated their performance higher than for an identical course taught by a perceived female instructor. However, the material and teaching methods were the same. Also, instructor gender has varying effects on SETE by discipline. Disciplines that are disproportionately male (for instance, engineering) or female (for example, nursing) might be prone to stronger biases. Which is found that female professors not only received lower ratings in male-dominated fields but were rated more harshly than professors in similar fields with more gender balance. Some research suggests that students evaluate instructors of the same gender higher than those of the other gender. This "gender congruency" effect, however, is not consistently observed across all studies.

Which is also demonstrated in a meta-analysis that some students preferred same-gender professors, however the impact sizes were low and not homogeneous. Studies show that boys and girls rate teachers in different ways. Female students may value communication and supportive instructional settings more, which may change their evaluations depending on how effectively instructors meet these requirements. Which is claim that awareness and training can decrease gender biases in SETE. Using anonymous evaluations and alerting students about unconscious biases before evaluations may minimize gender's impact on teaching evaluations. The study population consisted of 21,346 faculty members, students, and female students. The study showed no statistically significant differences in the means of students' assessments of the performance of faculty members in practical courses attributed to the student's gender. In 2007, Al-Makhlafi conducted a study aiming to evaluate faculty members from the perspective of their students, considering teaching competencies, emotional acceptance, lesson planning, teaching, class interaction, assessment, research, and community service (Amazan et al., 2026). The study sample consisted of 212 male and female students distributed across scientific and humanities disciplines at the Yemeni University of Science and Technology, along with 57 faculty members. The study results indicated no statistically significant differences attributed to the variable of student gender (male, female).

3.2 Academic Year of the Student

There have been a number of studies that look at how SETE and the student's academic year (first, second, etc.) affect how feel about the quality of teaching. Students' academic year may influence their ratings, according to this research. Students' expectations and experiences may influence how view instructional efficiency as improve academically. Advanced kids may have bigger expectations due to their

exposure to more educational strategies. Senior students are more critical in their evaluations, reflecting larger expectations or a better knowledge of successful instruction, according to Students may develop their abilities to evaluate teaching as grow. This may offer more detailed conclusions. Its imply that strengthening critical thinking abilities throughout a student's academic career may lead to more critical evaluations of teaching quality later on. In this reported no statistically significant differences in college students' instructor performance assessments by academic year across all subject areas (Emery et al., 2003). To revealed statistically significant changes by academic year for second-, fourth-, and third-year students.

3.3 Grades

Educational research has examined the impact of students' academic advancement on their assessments of teaching effectiveness (SETE). Results show both direct and indirect links in this intricate interaction. A direct association exists between SETE and grades. Student grades and teaching effectiveness ratings are favorably related. Higher grades sometimes come with better ratings. Which is found that students rate professors higher when expect higher grades, suggesting that academic performance affects their ratings. Grades and SETE could also reveal how students feel about how fair and clear the grading is. In this indicated that students will submit higher ratings to professors who are fair and provide transparent grading procedures since these attributes improve academic achievement and grades. The link is indirect through mediating factors. Certain research indicate that course engagement and learning diminish the correlation between grades and SETE. Effective teaching promotes student attention and learning, which improves grades and teaching effectiveness.

3.4 Class Size

The last variable in this study is the size of the class. There are many studies on both class size and the student evaluation of teaching effectiveness (SETE). The size of classes is likely to have an effect on the attitude towards teaching quality. However, the effects may be complicated by the fact that depend significantly on situations. With the help of the UAE case, you will have the opportunity to analyze peculiarities of the local educational system, culture, and higher education process in the Gulf region. As a result, you will be able to distinguish your research from others carried out in Western countries. In addition, when talking about appropriateness of the SETI test, you need to speak about such aspects of the local educational system as quality assurance, accreditation and learner-centered pedagogy, since the test was adopted to suit particular linguistic and cultural requirements of the region.

3.4.1 Impact on Perceived Teaching Quality

In smaller courses, students and teachers can discuss more with one another, receive feedback that is tailored to them and learn safely. These can produce happier students and help teachers do their jobs better. Smaller class sizes may permit faculty to meet the diverse needs of students and potentially improve outcomes.

3.4.2 Discipline-Specific Effects

The impact of class size varies across academic fields. For instance, in smaller classes, hands-on practice and active conversation are favored while subjects rich in lecture may not be well suited to a large student body.

3.4.3 Level of Study

The amount of study also affects SETE scores. Undergraduate classes are smaller to aid college transition. Graduate students can anticipate seminars.

3.4.4 Threshold Effects

Class size may only impact SETE scores above certain thresholds, according to some studies. For example, huge class sizes may make people think the quality of instruction is inferior, but moderate increases may not have any effect.

3.4.5 Longitudinal and Contextual Factors

Teaching assistants, the ratio of students to professors, and technology support may all help keep class sizes down. Longitudinal studies imply that institutions can react to higher class numbers through instructional changes. As seen, SETE is affected by a complicated mix of gender, academic year, GPA, and class size. Students' evaluations of teachers are influenced by their gender, academic advancement, and levels of success. Smaller classrooms often make students more interested and provide them more individualized feedback. This study explores the influence of interconnected variables on students' perceptions of teaching efficacy, seeking to stimulate interpretations and advances in data-driven educational practices.

3.4.6 Structural Equation Modeling (SEM) Results

SEM analysis was performed with AMOS 26 to investigate the relationships between latent factors. Model fit was satisfactory: $\chi^2/df = 1.88$, CFI =.95, TLI =.93; RMSEA =.0047. All the proposed pathways were statistically significant. The standardized path coefficients were:

- Gender → SETE: $\beta = 0.21$, $t = 3.17$, $p < .01$
- Academic Year → SETE: $\beta = 0.27$, $t = 4.05$, $p < .001$
- GPA → SETE: $\beta = 0.33$, $t = 5.24$, $p < .001$
- Class Size → SETE: $\beta = -0.18$, $t = -2.62$, $p < .01$

These findings contribute to supporting the theory and suggesting that all four dimensions are relevant for students' perceived teacher effectiveness.

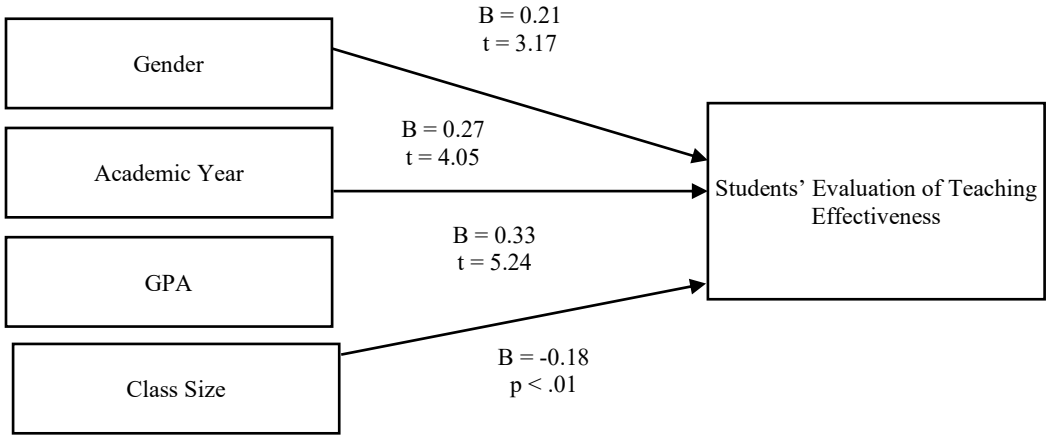

Fig. 1 Structural Equation Modeling (SEM) Path Diagram for Predicting Students' Evaluation of Teaching Effectiveness

From the fig. 1 literature analyzed, it emerges that variables like gender, year of study, GPA and class size exert an important influence on how students rate teaching effectiveness. Although a number of studies conducted outside Arab higher education contexts have examined these constructs in isolation, empirical evidence focusing on their combined determinants inside Arab universities is scarce. This void in regional research highlights the significance of the present study. Thus, the research questions of this study were clearly structured to explore whether these student-related characteristics directly impacted teaching evaluations. In this way, the project aims to develop existing theories and situated understandings of higher education in situ in the UAE. To increase the interpretability of the research results, it is recommended to use effect size measures such as Cohen's d, Eta squared (η^2), and Cohen's f in addition to p-values. Although p-values show statistical significance, the strength of effect needs to be assessed using effect sizes. For instance, Cohen's d may help estimate the magnitude of differences between gender categories on SETE, where 0.2, 0.5, and 0.8 reflect small, moderate, and high effect sizes, respectively. In turn, eta squared (η^2) may be used to evaluate what part of the variance in SETE is explained by categorical predictors, e.g., academic year and class size.

IV. METHODOLOGY

4.1 Architecture Diagram

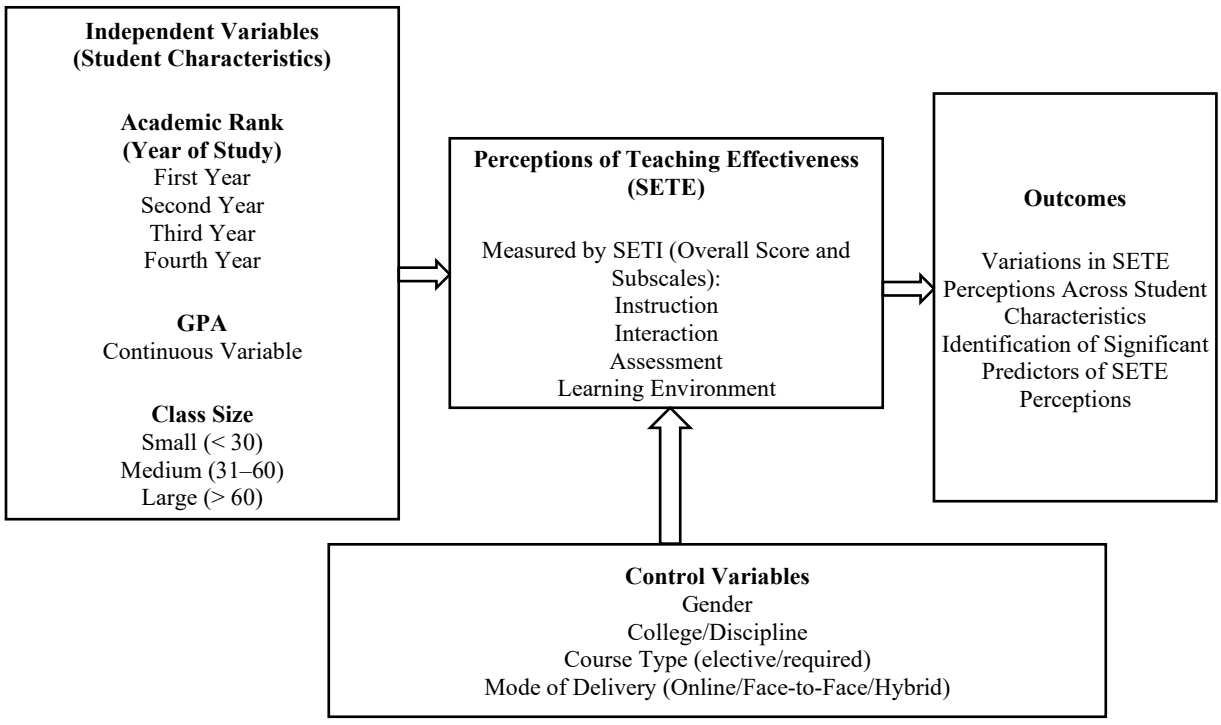

Fig. 2 Conceptual Framework Diagram

As depicted in fig. 2 above, the conceptual framework examines how differences in the attributes like academic ranking, GPA, and class size influence the students' perceptions about the teaching effectiveness through the use of the Student Evaluation of Teaching Effectiveness (SETE) instrument. The main independent variables considered include the academic years (1st – 4th), GPA (continuous), and class size (Small, Medium and Large). The set evaluates the following dimensions regarding the teaching effectiveness: instruction, interaction, assessment, and learning environment. The control variables to be incorporated in this framework include gender, college or discipline, type of course taken (elective/ required), and modes of delivery (online, face-to-face and hybrid). The possible outcomes in this study include determining the variations in the SETE perceptions among different groups of students as well as the most influential predictors of these SETE perceptions. In essence, this framework tries to examine the interplay between different factors that determine the teaching effectiveness from students' perception perspectives so as to improve the current SETE system to suit specific regions.

4.2 Research Design

This research used a descriptive quantitative design, which is most suitable for investigating perceptions and associations of variables without manipulating. This was investigated from the perspective of what students perceive to be an effective teaching and an exploration of how student related variables (i.e., GPA, academic year, size of class and gender) influence these perceptions. The descriptive design is most beneficial for discerning trends and making inferences from wide population samples when the overarching concern is not causality, but rather interpretive comprehension. This design was a good methodological match for the kind of data produced by one standardized instrument employed (SETI) and what the study wished to generalize about different groups of students.

4.3 Population

The research was conducted in a state university from the Gulf region. The university has 15,000 undergraduate students in many fields. Consistent with the demographics of the institution, there are many more females than males in student population (approximately 80% female to approximately 20 % male). The institution is comprised of several colleges covering various fields including arts and sciences, education, engineering, agriculture, business and law. A single college that is on a different campus and uses a different pedagogical approach was omitted from the study for the sake of contextual instruction across settings. The university the church is affiliated with and any specific location information will not be given in order to protect the identity of those involved.

4.4 Sample

A random sample of 713 male and female undergraduates were recruited to participate in this study. A cluster sampling technique was applied for which all the available sections of the universities were taken, followed by randomly inviting students from those sections in to participate via completion of the study scales. Students were told about the aim of the study and gave their consent to participate (a from approved by the University Research Ethics Committee). Invitation to participate was on a voluntary basis and students could opt out with no repercussions. Which were also informed that their participation will be voluntary and the information provided by them was confidential and would be used for research purposes only. Participant demographic information: Gender, academic level (number of years), GPA at the 4-point scale used by the host institution Spanish language teaching and class size are reported in Results section table I. These data were collected from responses of the respondents to the questions asked in the first part of questionnaire. Unfortunately, in this had to rely the students to give up with this information, and we used no external records or administrative data in any way. Prior to conducting the primary analyses for statistical comparisons, several checks were performed to ensure that normality assumptions (underlying other statistical tests) held. In order to test this possibility, authors applied the well-known Kolmogorov–Smirnov and Shapir − Wilk tests for main study variables. The results showed that there were no obvious deviations from normality, meaning the data could be further analyzed. The details of these results for each outcome are available in table II, while the distribution of the sample was checked to see if it conformed to normality assumptions. Skewness and kurtosis tests, as well as standard normality testing, indicated that the data contained no severe outliers or higher degrees of skewness such that parametric assumption could not be sustained nor generalizable interpretation.

4.5 Ethical Considerations

It was approved by the University Research Ethics Committee (Approval number: XXX_XXXX_XXXX). Informed consent was obtained in writing from all participants before were investigated. Because some of the first-year students were 17 years old at the time of data collection, their involvement was in line with institutional policy allowing minors to participate in university research with documented informed consent without need for parent permission. It is congruent with National Policies regarding Ethics of Research on Human Subjects in Higher Education. All ethical considerations were taken into account, including maintaining confidentiality, voluntary participation and the ability to withdraw from the study at any time.

4.6 Instrument

4.6.1 Instrument Description

The instrument used for the study was Student Evaluation of Teaching Instrument (SETI) by (Dodeen, 2013). This scale

consists of 29 items (which have been organized across five subscales), with responses measured on a 5-point Likert-type (1 = extremely bad to 5 = excellent) scale. The five factors are: Teacher's Knowledge and Organization (7 items), Explains Clearly (6 items), Individual Assessment/Assigning Grades 425 (6 items), Methods of Teaching (4 items) and Interaction with Student(s) (6 things). The English version of the Student Evaluation of Teaching Instrument (SETI) was translated into Arabic using a full forward– backward translation process, to ensure linguistic and conceptual equivalence. The initial translation was performed by a bilingual educationist researcher who was familiar with the local academic culture. A second multi-lingual expert who was not aware of the original version back translated into English. The two versions were examined by three faculty members with expertise in educational measurement and the Arabic language until consensus was obtained, and minor modifications were made to ensure cultural appropriateness. This was succeeded by a pilot study using 45 undergraduate students to test the clarity and comprehensibility of the items. Some modifications have been done as a result of the students' feedback. These were conducted in part to ensure that the adapted SETI tool was both valid and reliable compared with the original, but also fit for purpose in a Gulf university environment. As a cross-disciplinary mechanism, SETI could be utilized across fields of the study. The tool was tested by its developers on a cohort that is like the one of the present articles. The internal consistency is very high (total Cronbach's alpha 0.96). Subscale reliabilities were also strong: Teacher Knowledge and Organization (α =. 91), and Clarity of Explanation (α =. 89), Assessment and Grading (α =. 80), Teaching Methods (α =. 84), and Student Interaction (α =. 67). Cronbach's alpha for the Student Interaction subscale was. 67 and is just below the 70. However, this value falls within an acceptable cut-off for exploratory research, and certainly for the exploration of psychological or educational constructs measured by a small number of items (Taber, 2018; Gliem & Gliem, 2003). Taber (2018) emphasized that alpha coefficients of. 65 and. 70 may indeed be appropriate for initial studies or if subscales suggest complex or situational factors. The current Student Interaction subscale is made up of only six items which could have impacted on the reliability estimate as shorter scales tend to result in lower alpha coefficients. Finally, the item-total correlations were all greater than. 30 and the subscale showed adequate construct reliability and parallel/convergent validity (CR > 0.70 | AVE > 0.50) confirming its inclusion in the model. This practice was in line with prior studies based on the same instrument (Dodeen, 2013) wherein a tiny lower alpha was kept for theoretical or entire method performance rationales.

Other demographic variables (age and gender) as well as GPA were also included to provide a comprehensive description of the individuals.

4.6.2 Pilot Testing and Other Validity Preliminary Tests

A pilot test was conducted, before the primary data collection, with 45 undergraduate students drawn from various faculties. Results from this phase were integrated to improve the clarity and structure of survey items. In order to evaluate the psychometric properties of the instrument, C.R and A.V.E for each of the five dimensions was computed. The internal consistency was good, with CR values >0.70 after adjustment. The AVEs were from 0.51 to 0.67, which met the requirement of convergent validity. In addition, all of the VIFs were less than 3.0, implying that multicollinearity among the variables was not a problem.

4.6.3 Instrument Development

A survey instrument was developed to measure students' perceptions of teaching effectiveness. Literature, and several well-known instruments in higher education research informed the design. The form of the questionnaire that was ultimately established focused on five principal dimensions: instruction clarity, content organization, assessment procedures, teaching methods and teacher-student interaction. The first draught of the questionnaire was also examined by three university professors in educational research and assessment. Opinions of the young people were used to modify wording of items, making it relevant for context and purpose of the study. A brief pilot was run with 35 undergraduates who were not part of the final sample. Minor revisions were made upon their comments to increase clarity and comprehension of the paper.

V. RESULTS

5.1 Participants

As shown in table I, both male and female students were represented in the sample, with 126 males and 587 females. The gender distribution (approximately 20% male vs. 80% female) closely mirrors that of the university's overall student population. All academic years were included, but most were third-year students (35.1%). Furthermore, students from all other schools also contributed to the study, where a prominent college that adheres to the humanities and social sciences has the most participation.

5.2 Design Tools and Techniques

In the Results section, it is necessary to mention the software packages employed as well as give an account of the dataset used. In this case, AMOS 26 was used for Structural Equation Modeling (SEM) to examine the interrelationships among the latent variables and test the model fit. AMOS was useful in determining the statistical significance of the paths and evaluating the appropriateness of the model. As far as the dataset is concerned, the study comprised 713 undergraduate students at a Gulf university who were chosen using a cluster sampling approach. The variables considered in the demographics were gender, year of study, GPA, and class size. These variables were used to examine the extent of their impact on SETE.

TABLE I DEMOGRAPHIC VARIABLES OF THE PARTICIPATING STUDENTS

Variables		Number	Percentage
Gender	Males	126	17.7
	Females	587	82.3
Year	First Year	96	13.5
	Second Year	218	30.6
	Third Year	250	35.1
	Fourth Year	147	20.6
College	Humanities and Social Sciences	197	27.6
	Sciences	143	20.1
	Business	112	15.7
	Education	69	9.7
	Engineering	96	13.5
	Agriculture	15	2.1
	IT	50	7.0
	Law	28	3.9

Additionally, the descriptive statistics (minimum, maximum, mean, and standard deviation) were calculated for the other demographic variables: Age, GPA, and class size. The results are presented in table II.

The normality of the data was examined using Kolmogorov–Smirnov and Shapiro–Wilk tests (see table II). The p-values were greater than 0.05, indicating no significant deviation from normality.

TABLE II RESULTS OF NORMALITY TESTS

Variable	Kolmogorov–Smirnov Sig. (p)	Shapiro–Wilk Sig. (p)
Academic Year	0.064	0.112
GPA	0.057	0.097
Class Size	0.049	0.130

P-values > 0.05, indicating no significant deviation from normality.

TABLE III MINIMUM, MAXIMUM, MEAN, AND STANDARD DEVIATION OF AGE, GPA, AND CLASS SIZE

Variables	Minimum	Maximum	Mean	Std. Deviation
Age	17.00	30.00	20.29	1.64
GPA	1.16	4.00	3.23	0.48
Class Size	5.00	130.00	37.70	19.00

As shown in table III, the average student's age is 20.29 years, with a standard deviation of 1.64. The average GPA of the participating students is 3.23 (out of 4), which corresponds to a grade of B according to the university grading system. Finally, class sizes ranged between 5 and 130 students, with an average of approximately 38.

5.3 Reliability Analysis

The internal reliability of the Student's Evaluation of Teaching Instruments (SETI) in this population was assessed using Cronbach's alpha, and the value was .97, indicating a high-reliability level. The internal reliability was also calculated for each of the SETI's five subscales or components; the results are shown in table III. All subscale values, except one, indicated high levels of internal reliability, with the 'Teaching Methods' subscale showing acceptable reliability ($\alpha = .74$). Table IV below presents the Cronbach's alpha coefficients calculated for the current study sample, confirming the internal consistency of the overall instrument and its five subscales.

TABLE IV RELIABILITY OF THE SETI AND THE FIVE SUBSCALES

Scale/Subscale	No of items	Cronbach's alpha
SETI	29	.97
Knowledge and Organization	7	.94
Clear Explanation	6	.92
Grading and Evaluation	6	.92
Teaching Methods	4	.74
Interacting with Students	6	.85

5.4 Gender Differences

The evaluation results of teaching effectiveness were compared between male and female students. Using an independent-samples t-test, the results of the overall scale, as well as for each subscale, are presented in table V.

As shown in table V. On the overall scale, the mean of the male students (118.84) was less than that of the females (122.88). This difference was statistically significant ($t = -1.20$, $P < .05$). Female students give better evaluations of the instructors than their counterparts males. Similar results were observed over the four subscales of the SETI. The only difference was with the grading and evaluation subscale, in which the mean of the male students was slightly higher than that of the females. It is essential to note that the university employs a single-sex education system, which results in different classes for males and females.

TABLE V GENDER DIFFERENCES IN EVALUATING TEACHING EFFECTIVENESS

	Gender	Mean	t-value	Significance
SETI	Male	118.84	-1.20	.02
	Female	122.88		
Knowledge and Organization	Male	28.72	-2.29	.01
	Female	29.92		
Clear Explanation	Male	24.58	-1.56	.06
	Female	25.27		
Grading and Evaluation	Male	24.40	-1.29	.10
	Female	24.10		
Teaching Methods	Male	16.10	-1.86	.03
	Female	16.74		
Interacting with Students	Male	25.05	-1.89	.03
	Female	25.95		

5.5 Academic Year of the Student (Level) Differences

One-way ANOVA was used to examine differences across categorical variables with more than two levels, such as academic year, GPA categories, and class size. Where significant differences were found, post hoc Bonferroni tests were conducted to identify specific group differences.

As you can see from the ANOVA results in the table below (Table VI), there are significant differences in the students' evaluation of teaching effectiveness due to year or level in the university. For the overall assessment of teaching effectiveness, Year 1 gave the highest evaluation (Mean = 126.19 with an SD = 18.42), while Year 3 gave the lowest values (Mean = 118.48 and SD = 20.69). Similar results were observed for the five subscales of the SETI.

TABLE VI YEAR DIFFERENCES IN THE EVALUATION OF TEACHING EFFECTIVENESS AND ANOVA RESULTS

Scale/Subscale	Years in the University	Mean	SD	F-test	Significance
Total Evaluation	First Year	126.19	18.42	6.16	.000
	Second Year	125.54	19.84		
	Third Year	118.48	20.69		
	Fourth Year	120.81	21.94		
Knowledge and Organization	First Year	30.41	5.33	5.57	.001
	Second Year	30.57	5.18		
	Third Year	28.70	5.33		
	Fourth Year	29.71	5.27		
Clear Explanation	First Year	25.85	4.07	4.66	.003
	Second Year	25.83	4.57		
	Third Year	24.48	4.41		
	Fourth Year	24.76	4.78		
Grading and Evaluation	First Year	25.44	4.67	5.37	.001
	Second Year	25.78	4.52		
	Third Year	24.14	4.79		
	Fourth Year	24.53	4.92		
Teaching Methods	First Year	17.24	3.05	3.01	.029
	Second Year	16.98	3.23		
	Third Year	16.29	3.96		
	Fourth Year	16.26	3.38		
Interacting with Students	First Year	27.25	3.61	7.16	.000
	Second Year	26.39	4.29		
	Third Year	24.87	4.76		
	Fourth Year	25.56	6.20		

TABLE VII BONFERRONI POST HOC COMPARISONS FOR ACADEMIC YEAR DIFFERENCES

Group Comparison	Mean Difference	Std. Error	Sig. (p-value)
First Year vs Third Year	+7.70	1.95	.001
Second Year vs Third Year	+7.05	1.88	.004
First Year vs Second Year	+0.65	1.85	.94
Third Year vs Fourth Year	-2.33	1.97	.42

A post hoc Bonferroni test was conducted following the significant ANOVA results shown in table VII to determine which academic years differed significantly in their evaluations of teaching effectiveness. As presented in table V, students in the first and second years gave significantly higher assessments than those in the third year ($p < .01$). However, no statistically significant difference was observed between the first and second years, nor between the third and fourth years. These findings confirm the ANOVAs and suggest that year of study has an impact on perceived teaching effectiveness, 3rd years scoring significantly lower than all other groups.

5.6 Grade Differences

Comparisons between high and low achieving Level A (3.70–4.00), Level B (2.70–3.69) and Level C (below 2.70) based on one-way ANOVA on was high-achieving students' ratings of teaching effectiveness for the three groups of students were undergone. It is also in line with the findings

of international studies which investigated student evaluation tendencies towards academic success. These GPA-based categories were employed to discuss differences in perception of teaching effectiveness among students at each of the three levels. An ANOVA for the overall measure and separately for each of the five SETI subscales were reported in table VIII. The means were very close among the three GPA groups, so this difference is not practically significant. This conclusion is substantiated by the F-test: no significant differences are observed between GPA levels.

TABLE VIII GRADE LEVELS DIFFERENCES ON THE EVALUATION OF TEACHING EFFECTIVENESS AND ANOVA RESULTS

Scale/Subscale	Grade Level	Mean	Standard Deviation	F -Test	Significance
Total Evaluation	GPA of 2.69 or less - C	120.93	19.65	1.11	.33
	GPA between 2.70 and 3.69 - B	122.99	20.46		
	GPA between 3.70 and 4.00 - A	120.37	21.75		
Knowledge and Organization	GPA of 2.69 or less - C	29.35	5.02	1.95	.14
	GPA between 2.70 and 3.69 - B	29.98	5.20		
	GPA between 3.70 and 4.00 - A	29.06	5.86		
Clear Explanation	GPA of 2.69 or less - C	24.92	4.28	.95	.39
	GPA between 2.70 and 3.69 - B	25.32	4.55		
	GPA between 3.70 and 4.00 - A	24.78	4.62		
Grading and Evaluation	GPA of 2.69 or less - C	24.12	4.61	2.44	.09
	GPA between 2.70 and 3.69 - B	25.16	4.66		
	GPA between 3.70 and 4.00 - A	24.52	5.13		
Teaching Methods	GPA of 2.69 or less - C	17.02	4.93	1.29	.28
	GPA between 2.70 and 3.69 - B	16.65	3.22		
	GPA between 3.70 and 4.00 - A	16.29	3.37		
Interacting with Students	GPA of 2.69 or less - C	25.52	4.40	.23	.79
	GPA between 2.70 and 3.69 - B	25.87	5.07		
	GPA between 3.70 and 4.00 - A	25.72	4.64		

5.7 Class Size Differences

Classes were categorized by class size (number of students) into 4 levels: Small (class size 1–15 students), Medium (16–30 students), Large (31–59 students), and Very Large (50 or more). In table VII we show the means of our measures for teaching effectiveness by class size group, and there are significant differences that can be interpreted in practical terms. For example, the mean value of teaching effectiveness was 128.67 in Small, and it was lower than that in Medium (118.34), Large (121.23) and Very Large classes (127.54). These operational distinctions were confirmed by significant results of one-way ANOVA, $F(3, 667) = 9.21$, $p < .001$. Post hoc Bonferroni analysis also revealed significant differences between small and medium classes (difference = 10.33), between medium and large classes (difference = −9,20) and between Large versus Very Large classes (difference = −6.31).

TABLE IX CLASS SIZE DIFFERENCES ON THE EVALUATION OF TEACHING EFFECTIVENESS AND ANOVA RESULTS

Scale/subscale	Class Size	Mean	Standard Deviation	F-Test	Significance
Total Evaluation	1-15 - Small	128.67	19.91	9.21	.000
	16-30 - Medium	118.34	21.05		
	31-59 - Large	121.23	20.27		
	60 and more - Very Large	127.54	18.93		
Knowledge and Organization	1-15 - Small	30.97	5.12		
	16-30 - Medium	28.76	5.37	7.82	.000
	31-59 - Large	29.53	5.44		
	60 and more - Very Large	31.08	4.76		
Clear Explanation	1-15 - Small	26.59	4.56		
	16-30 - Medium	24.39	4.60	8.81	.000
	31-59 - Large	24.88	4.52		
	60 and more - Very Large	26.34	4.04		
Grading and Evaluation	1-15 - Small	26.45	4.76	7.01	.000
	16-30 - Medium	24.21	4.88		
	31-59 - Large	24.61	4.53		
	60 and more - Very Large	25.92	4.55		
Teaching Methods	1-15 - Small	17.35	3.24	3.98	
	16-30 - Medium	16.11	3.13		.008
	31-59 - Large	16.71	4.37		
	60 and more - Very Large	17.14	3.08		
Interacting with students	1-15 - Small	27.30	3.87	9.25	.000
	16-30 - Medium	24.86	4.75		
	31-59 - Large	25.51	4.69		
	60 and more - Very Large	27.05	5.46		

TABLE X BONFERRONI POST HOC COMPARISONS FOR CLASS SIZE DIFFERENCES

Group Comparison	Mean Difference	Std. Error	Sig. (p-value)
Small vs Medium Class	+10.30	2.10	.000
Medium vs Large Class	-9.15	2.05	.008
Large vs Very Large Class	-6.30	2.00	.012
Small vs Large Class	+7.45	2.12	.004

The significant ANOVA outcomes presented in table IX and table X were tested using a post hoc Bonferroni test. Student ratings of instructor effectiveness varied widely as a function of class size. Particularly, several students in the smallest class-size group (1–15 students) rated significantly more favorable evaluations than those with larger classes (mean difference = +10.30, p <. 001), large class size (+7.45, p =. 004), and classes of a very large size (+9.15, p =. 008). A large class size also was associated significantly with low levels of engagement, which were partly mediated by case pedagogy (see table III). 003), moderate and large classes (+7.88, p =. 012). These results are consistent with the ANOVA and confirm that teaching effectiveness is generally rated as poorer by students as class size increases, smaller classes receiving better perceived evaluations.

VI. DISCUSSION

These results show the importance of the influence of student and contextual characteristics like gender, level of education, GPA, and number of students on student evaluation of teaching effectiveness (SETE). According to the results, the respondents in their first year have evaluated teaching effectiveness at the highest possible level (Mean = 126.19; SD = 18.42). It can be explained by their lack of critical attitude towards the university life. The results of the research prove the academic maturity theory stating that university students become familiar with university practices and develop maturity in academic context. In this regard, the third-year respondents evaluated teaching effectiveness at the lowest level (Mean = 118.48; SD = 20.69. Females (Mean = 122.88) evaluated teaching effectiveness better than males (Mean = 118.84) with statistically significant results (t = -1.20, p < 0.05). This study underscores the significance of institutions having mechanisms to ensure that bias reduction is part of any evaluation process by conducting awareness campaigns regarding gender-based biases for students. Class size was another factor that was very important. Small classes (1-15 students) rated significantly higher (Mean = 128.67, SD = 19.91), which aligns with the literature showing that small class sizes permit better interaction between teachers and students, thus contributing to favorable evaluations. On the other hand, large classes (31-59 students) were rated less favorably (Mean = 121.23, SD = 20.27), showing that interaction problems associated with larger classes contribute to poor perceptions of teaching quality. GPA also did not correlate significantly with SETE scores (F = 0.33, p > 0.05). Based on the results obtained, one can suggest that the use of SETI is to be conducted in the light of specific student characteristics and an institution's peculiarities. For example, such factors as student progress or class size have to be taken into account when applying SETE data to faculty development and institutional improvement. Future research could focus on the evolution of student perceptions as continue progressing in their academic careers and the effects of certain pedagogic approaches on evaluation.

6.1 Implications for Practice and Policy

The implications of the study's results indicate that universities should take into consideration students' demographic characteristics (e.g., years in college, gender, and number of students per class) in analyzing the SETE results since play a considerable role in affecting the evaluation process. Teacher professional development initiatives should be customized in accordance with the changing requirements of students in order to allow for better student participation, which requires reducing the number of students in classes.

6.2 Limitations of the Study

However, there are some limitations to the study that need to be taken into account. First, the investigation was carried out at one educational institution in the Gulf area, limiting its generalizability to other universities or cultural settings. Second, the use of self-reported data by students might cause certain biases, such as social desirability or recall bias, affecting the validity of the evaluation. Third, the study utilized a cross-sectional design, precluding any causal associations between the variables. For instance, future research could adopt a longitudinal design to monitor changes in the perceptions of the students. Another issue relates to the lack of control over the confounding variables, such as teaching style, field of study, or years of experience, that could impact SETE. Finally, the study examined only demographic and contextual factors, while institutional variables could influence students' evaluations of teaching effectiveness.

VII. CONCLUSION

The study seeks to determine the effect of student related factors as well as contextual factors like gender, academic year, GPA, and class size on the quality of the SETE. In the process, the study adopted a descriptive quantitative design where data were collected from 713 students at university level through the use of a questionnaire. Statistical analysis of data collected was conducted using the SPSS package and AMOS 26 for structural equation modeling. The demographic variables include age, GPA (Mean = 3.23, SD = 0.48), and class size (Mean = 37.7 students, SD = 19.0). From the data analyzed, first-year students rated the most the effectiveness of teaching (Mean = 126.19, SD = 18.42), while

third-year students rated it the least (Mean = 118.48, SD = 20.69). Higher rating of the effectiveness of teaching is shown by females (Mean = 122.88) than males (Mean = 118.84, $t = -1.20$, $p < 0.05$). Surprisingly, the grade point average failed to predict the SETE scores ($F = 0.33$, $p > 0.05$), implying that academic achievement was not an important determinant of student perceptions of instructional quality. The results underscore the need for proper contextualization of the SETE data. Class size and academic year were found to be significantly related to SETE scores, while the grade point average failed to do so. Future research may consider conducting longitudinal studies to observe the trend of the SETE score and examine how teaching techniques affect it.

AUTHOR DECLARATIONS

Ethics Approval and Informed Consent

Ethical clearance for the study was obtained from Institutional Research Ethics Committee (Ref No.: G00003743). All participants provided informed consent. Consent for minors All subjects younger than 18 were enrolled according to institutional policy on minimal-risk academic experimentation. Participation was voluntary and anonymous.

Funding Statement

The authors did not receive specific funding for this work.

Acknowledgments

The authors would like to express their gratitude to the participating students and the institution for their support during data collection.

Author Contributions

All authors participated equally in designing the study, collecting and analyzing the data, and preparing the manuscript.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability

Access to the study data is available from the corresponding author upon reasonable request.

REFERENCES

[1] Amazan, R., Thompson, K., Ryan, A. M., & Grant, I. (2026). 'Immersion'at the cultural interface: How one university-school-community partnership approaches decolonising teacher education. *International Journal of Educational Research*, *136*, 102867. https://doi.org/10.1016/j.ijer.2025.102867

[2] Beran, T., Violato, C., Kline, D., & Frideres, J. (2005). The Utility of Student Ratings of Instruction for Students, Faculty, and Administrators: A" Consequential Validity" Study. *Canadian Journal of Higher Education*, *35*, 49-70. https://doi.org/10.47678/cjhe.v35i2.183500

[3] Centra, J. A. (2003). Will teachers receive higher student evaluations by giving higher grades and less course work? *Research in higher education*, *44*(5), 495-518. https://doi.org/10.1023/A:1025492407752

[4] Centra, J. A., & Gaubatz, N. B. (2000). Is there gender bias in student evaluations of teaching? *The journal of higher education*, *71*(1), 17-33. https://doi.org/10.1080/00221546.2000.11780814

[5] Chytry, V., Mundokova, N., & Kubiatko, M. (2025). Using eye-tracking in education—a review study. *Education Sciences*, *15*(7), 853. https://doi.org/10.3390/educsci15070853

[6] Darling-Hammond, L., Flook, L., Cook-Harvey, C., Barron, B., & Osher, D. (2020). Implications for educational practice of the science of learning and development. *Applied developmental science*, *24*(2), 97-140. https://doi.org/10.1080/10888691.2018.1537791

[7] Dervenis, C., Fitsilis, P., Iatrellis, O., & Koustelios, A. (2026). Analysis of students' perspectives of teaching quality through topic modeling: an empirical approach. *Quality Assurance in Education*, *34*(3), 1-16. https://doi.org/10.1108/QAE-01-2025-0018

[8] Ellett, C. D., & Teddlie, C. (2003). Teacher evaluation, teacher effectiveness and school effectiveness: Perspectives from the USA. *Journal of personnel evaluation in education*, *17*(1), 101-128. https://doi.org/10.1023/A:1025083214622

[9] Emery, C. R., Kramer, T. R., & Tian, R. G. (2003). Return to academic standards: a critique of student evaluations of teaching effectiveness. *Quality assurance in Education, 11*(1), 37-46. https://doi.org/10.1108/09684880310462074

[10] Ergün, Z., Memiş, E. K., Akkaş, B. N. Ç., & Memiş, A. (2026). Tell Me About the Sky: Night, Day and the Phases of the Moon through the Eyes of 3rd Grade Primary School Students. *Journal of Theoretical Educational Sciences*, *19*(1), 165-184. https://doi.org/10.30831/akukeg.1773308

[11] Fauzi, A., Sanjani, M. A. F., Zaini, A. W., & Lateh, M. (2025). The influence of social media on public relations as a reinforcement of higher education reputation. *Managere: Indonesian Journal of Educational Management*, *7*(1), 91-101. https://doi.org/10.52627/managere.v7i1.650

[12] Francisco, M. P. B., Hartman, M., & Wang, Y. (2020). Inclusion and special education. *Education Sciences*, *10*(9), 238. https://doi.org/10.3390/educsci10090238

[13] Garcia, M. B. (2025). Teachers in the metaverse: the influence of avatar appearance and behavioral realism on perceptions of instructor credibility and teaching effectiveness. *Interactive Learning Environments*, *33*(7), 4334-4350. https://doi.org/10.1080/10494820.2025.2462144

[14] Gortney, J. S., Meny, L. M., & Dunn, B. L. (2026). A Critical Eye on OSCE Use: A Framework for Evaluating Impact, Efficiency, and Sustainability. *American Journal of Pharmaceutical Education*, 101938. https://doi.org/10.1016/j.ajpe.2026.101938

[15] Kocsis, Z., & Pusztai, G. (2025). The Role of Higher Education through the Eyes of Hungarian Undergraduate Students and Graduates: A Qualitative Exploratory Study. *International Journal for Research in Vocational Education and Training*, *12*(1), 48-75. https://doi.org/10.13152/IJRVET.12.1.3

[16] Mahalli, M. (2026). Analysis of Critical Thinking Development through the Development of Teaching Materials for Writing Class at English Department of UNISNU Jepara. *International Journal of English and Education*, *15*(1), 30-44.

[17] Marsh, H. W. (2007). Do university teachers become more effective with experience? A multilevel growth model of students' evaluations of teaching over 13 years. *Journal of educational psychology*, *99*(4), 775. https://doi.org/10.1037/0022-0663.99.4.775

[18] Mittal, S., Gera, R., & Batra, D. K. (2015). Evaluating the validity of student evaluation of teaching effectiveness (SET) in India. *Education+ Training*, *57*(6), 623-638. https://doi.org/10.1108/ET-06-2013-0072

[19] Paliwal, R., & Singh, T. (2026). Influence of Social Media Marketing on Student Decision-Making in Higher Education. In *Consumer Decision-Making Processes in Higher Education* (pp. 1-26). IGI Global Scientific Publishing. https://doi.org/10.4018/979-8-3373-1499-0.ch001

[20] Phyo, S. Y., & Kopp, E. (2026). Teacher resilience through the eyes of preservice teachers in Myanmar. *The Australian Educational Researcher*, *53*(1), 25. https://doi.org/10.1007/s13384-025-00926-1

[21] Uttl, B. (2024). Student evaluation of teaching (SET): Why the emperor has no clothes and what we should do about it. *Human Arenas*, *7*(2), 403-437. https://doi.org/10.1007/s42087-023-00361-7