

## Validating Student Evaluations of Teaching in Teacher Education Programs at the University of Sargodha

Farkhanda Siddiqui

[farkhandasiddiqui548@gmail.com](mailto:farkhandasiddiqui548@gmail.com)

Ph.D. scholar, Institute of Education, University of Sargodha, Punjab, Pakistan

Dr. Abdul Majeed Khan Niazi

[abdul.majeed@uos.edu.com](mailto:abdul.majeed@uos.edu.com)

Assistant professor, Institute of Education, University of Sargodha, Punjab, Pakistan

Farhana Mushtaq

[farhanamushtaq7402@gmail.com](mailto:farhanamushtaq7402@gmail.com)

Ph.D. scholar, Institute of Education, University of Sargodha, Punjab, Pakistan

Corresponding Author: Farkhanda Siddiqui [farkhandasiddiqui548@gmail.com](mailto:farkhandasiddiqui548@gmail.com)

Received: 25-01-2026

Revised: 08-02-2026

Accepted: 23-02-2026

Published: 10-03-2026

### ABSTRACT

*Teaching quality is an important part of learning and engagement in higher education and a major factor in the success or failure of individual students. Student Evaluations of Teaching (SETs) have become a common measure to gauge teaching effectiveness and guide instruction. This study confirmed the SETs at University of Sargodha and compared the students' perception based on gender, program level and semester.*

*The method used was a quantitative cross-sectional design. A questionnaire adapted from the SET instrument which included questions on planning and teaching didactics, and assessment and feedback was administered to 170 BS, MPhil and PhD students at the end of the semester. The instrument was validated using reliability, confirmatory factor analysis, independent-samples *t* test, ANOVA, correlation and regression analyses and patterns of perception were explored.*

*The instrument has high reliability (Cronbach's alpha = .92) and construct validity. Pupils' overall views of the quality of teaching were positive. Female students gave higher ratings than male students in the categories of planning, didactics and assessment. There were differences between the perceptions of academic stages with BS 4th semester students showing higher satisfaction on planning than those of BS 8th semester students, while PhD students' perceptions were higher on assessment and feedback than MPhil students' perceptions. There was a high level of support between planning and teaching didactics and assessment and feedback ( $R^2 = .681$ ). SETs were found to be a reliable instrument for instructional improvement, and demographic and academic differences were used to help determine the appropriate strategies.*

**Keywords:** *Student Evaluations of Teaching, teacher education, teaching effectiveness, assessment and feedback, instructional planning, University of Sargodha, construct validity, higher education.*

### INTRODUCTION

In higher education, the effectiveness of teaching is recognized as an important factor to ensure student learning outcomes, engagement, and long-term success in studies (Quansah et al., 2024). There are several methods used by universities around the world to guarantee the quality of instruction, one of which is called student evaluations of teaching (SETs) and is widely used and one of the most common

(Cook et al., 2022). SETs are designed to offer instant feedback on how well the instructor performs the lesson, including various aspects like planning for the lesson, clarity in presenting material, engagement strategies, assessment methods, and classroom interaction (Daumiller et al., 2022). The evaluations provide institutions with actionable data which can drive improvement efforts in teaching, curriculum development and faculty professional development, leading to better student learning experiences.

Although widely used, SETs have a number of limitations. New studies emphasize issues with reliability, validity and fairness of SETs, especially as the only measure of faculty performance (Lakeman et al., 2022). Student bias, difficulty of courses and instructors, or other factors unrelated to teaching quality can also systematically influence evaluations and result in misjudgments of teaching effectiveness (Heffernan et al., 2022). It is therefore crucial to develop valid and reliable tools for measuring the multidimensionality of teaching activities with minimal amount of extraneous bias. To ensure the validity of measurements, the factor structure of assessment instruments is examined, internal consistency is tested, and measurement invariance is validated for different groups of students (academic stage, gender, program level) (Hooda et al., 2022). However, if this is not done, differences in observed scores for the SET could be due to cohort variations, not to differences in the quality of the teaching.

The planning and didactic methodology is one of the important dimensions of teaching effectiveness, which refers to the teacher's preparation plans, the clarity of teaching goals, the congruence of what the teacher teaches with the previous learning, and using learning strategies that support students' understanding and interest (Khlaif et al., 2024). The planning process helps to organize the learning space and helps to facilitate meaningful learning through interactive, collaborative approaches (Berk & Theall, 2023). Likewise, another dimension is assessment and feedback, as there is a need to encourage students to learn continuously, which can be effectively achieved with timely, constructive, and transparent feedback (Hooda et al., 2022). Assessment practices, such as formative and summative feedback, have been found to facilitate students developing cognitive skills and socio-emotional skills, as well as recognizing strengths and opportunities for improvement (Maki & Kuh, 2023).

The use of digital tools and blended learning spaces creates additional challenges in the field of assessment of teaching. While traditional SET instruments cannot completely cover the capability of teachers to provide students with virtual learning, use digital resources effectively and encourage students to be collaborative in virtual learning (Almulla et al., 2023). The assessment of the SETs should address modern pedagogical approaches, digital skills, and digital literacy because in higher education the blended and technology-mediated teaching dominant method requires the integration of these elements in the validation process (Khassawneh et al., 2022). This gives a comprehensive picture of teaching effectiveness in both the physical and digital classroom setting to inform decisions about improving teaching and learning.

The University of Sargodha has a system of regular SETs in the Department of Education to collect feedback from students on the teaching performance of teachers. But the psychometric aspects of these instruments have not been extensively studied and so there are some concerns about validity, reliability, and fairness. In particular, it remains unknown whether current SETs are measuring the desired constructs (planning, teaching strategies and assessment practices) or whether the results are being affected by student bias, the semester or the level of the program (Okoye & Arrona-Palacios, 2022). If the differences in scores between groups, for instance, 4th semester BS and 8th semester BS students, or 2nd semester MPhil and PhD students, are not validated, then they might only be the result of different measurements and not of any differences in the effectiveness of the instruction given by teachers (Constantinou & Wijnen-Meijer, 2022).

Furthermore, SETs may not be comprehensive enough in documenting innovative pedagogical practices, such as those involving interaction learning, digital learning platforms, and socio-emotional learning, among others. When measures are inadequate, they can lead to inaccurate data, which can negatively impact teacher evaluation and hinder the use of data to guide action on teacher practice (Khlaif et al., 2024). The evidence of measurement invariance is also limited, which is necessary to make sure that the SETs work equally for various student groups. In the absence of invariance testing, assumptions based on group comparisons may be false, which could affect the advantages or disadvantages of certain instructors or courses based on the composition of the student groups they teach (Daumiller et al., 2022).

Effective teaching is an important component of higher education, directly affecting student learning, engagement and professional readiness. Limited empirical evidence exists in the Education Department, University of Sargodha, about students perception towards teaching performance in various program levels and different semesters of the education department. The lack of this knowledge can present a difficulty for administrators and faculty to determine where instructional improvements are needed or where specific professional development is needed.

In addition, contemporary higher education focuses on cultivating critical thinking, team learning and socio-emotional competencies, but the effectiveness of these competencies in the delivery and receipt of higher education has not been established for varying student groups. To understand the patterns of the teaching effectiveness, as perceived by prospective teachers in the terminal semester is very pivotal, as it will enable curriculum design and development, the improvement of pedagogical strategies and the proper alignment of teaching practices with the local educational objectives and international standards. Thus, the purpose of this study was to examine student perceptions of the level of instruction and student perceptions of differences between instructional program levels and semesters and to offer insights to improve instruction.

## **LITERATURE REVIEW**

Student evaluations of teaching (SETs) are considered a very important instrument to measure the effectiveness of teaching and to improve instruction in higher education (Quansah et al., 2024). These assessments offer feedback on various aspects of teaching, such as planning, the organisation of teaching and learning, delivery, assessment methods, and engagement in the classroom, and can help institutions to track the quality of teaching and improve student learning (Cook et al., 2022). SETs are often used for pedagogical development as well as faculty assessment, promotion and curriculum planning. Yet, there is a consistent debate of the reliability and validity of SETs, especially when using them as the sole source for making faculty evaluation decisions (Daumiller et al., 2022). Recent research highlights that non-teaching related factors, including difficulty of the course, personality of the teacher and the gender of the teacher, can impact students' perceptions and that this must be taken into account for fair and accurate assessment (Lakeman et al., 2022).

In the current study, the role of the psychometrically valid instruments has been emphasized, as instruments that are not valid can give false results and affect negatively the teacher and the student (Hooda et al., 2022). Others, such as Berk et al., (2023) suggest that SEs might not be able to distinguish between effective and ineffective teaching and could also have biases embedded within student perceptions of the teaching, rather than the quality of the instruction. Heffernan et al. (2022) also highlight the importance of using instruments that reduce systematic bias and are consistent across different cohorts of students. In this context, validation refers to an analysis of the internal consistency, factor structure and measurement invariance of the SET instruments, so the instrument reflects the intended dimensions of teaching performance across different types of students such as undergraduate and postgraduate students (AlAli & Al-Barakat, 2022).

It is well established that teaching effectiveness is multidimensional, comprising of planning and didactic strategies, assessment and feedback, and student engagement (Khlaif et al., 2024). Planning and didactics include the structuring of the course, precise expression of objectives, connection with what was already learned, strategies used for meaningful learning. Assessment and feedback includes the use of fair, transparent and timely forms of assessment to foster ongoing learning and development for students (Maki & Kuh, 2023). Engagement is the term used to describe what instructors do to motivate students, encourage collaboration and build socio-emotional skills, all of which are now acknowledged as essential to higher education success (Almulla, 2023). These dimensions offer a comprehensive overview of teaching quality and a basis for creating instruments that will yield valid, reliable and actionable feedback.

Recent studies have identified implementation of digital tools and blended teaching and learning environments as factors that have complicated the assessment of teaching effectiveness (Khassawneh et al., 2022). Older digital SETs may not adequately reflect the abilities needed to integrate students into online or blended classes, such as how to use learning management systems, facilitate collaborative online learning, and incorporate interactive content. Research indicates that incorporating these skills into the SET instruments has the potential to enhance evaluation relevance and applicability, and to offer more comprehensive insights into instructional quality (Berk & Theall, 2023). Additionally, the effective use of technology in instruction has been associated with a higher level of student engagement, motivation and achievement (AlAli & Al-Barakat, 2022), highlighting the need to assess these skills.

Many studies have focused on the correlation between SETs and students' learning outcomes, with findings indicating that increased teaching effectiveness (as measured by validated SET instruments) correlates with better cognitive and socio-emotional outcomes (Xia et al., 2024). Highly effective instructors are better able to foster critical thinking, problem solving, and cooperative learning which ultimately leads to improved student learning. Furthermore, validated SETs deliver feedback for professional development and targeted interventions and allow institutions to implement measures that enhance the quality of teaching and student satisfaction (Constantinou & Wijnen-Meijer, 2022). The evidence shows that when psychometrically sound, SETs can be used as a diagnostic and developmental tool to span the gap between assessment and instructional improvement.

A growing awareness of the importance of measurement invariance is a methodological concern in the study of SETs (Okoye & Arrona-Palacios, 2022). Measurement invariance means the instrument will produce equivalent measurements within each group (e.g., amongst students at different program levels, in different semester, or different gender groups). Comparisons between groups may lead to erroneous conclusions without the establishment of invariance. Recent empirical research, on the other hand, shows that perceptions of teaching effectiveness can be systematically different based on academic stage, program type, or prior experience in the program, highlighting the need to confirm the validity of SETs when considering them across groups (Daumiller et al., 2022; Heffernan et al., 2022). The assumption of measurement invariance is thus a prerequisite for meaningful analysis of differences in teaching performance and is an important component in equitable assessment.

There has been growing research on novel pedagogical approaches, such as fostering critical thinking, joint learning, and socio-emotional skills in HE (Almulla, 2023; Khlaif et al., 2024). The instruments used in SET that include items measuring these competencies are more likely to give feedback reflecting the current priorities of education. Further, research indicates that the dimensions found in SETs are consistent with quality standards and accreditation criteria across the world, which allows for assessments to be linked to wider institutional aims in relation to student learning and development (Hooda et al., 2022). Institutions can use these measures to gain a more comprehensive understanding of teaching

effectiveness than is provided by conventional measures, such as how well students understand the material or the speed at which they receive grades.

Faculty development is also an important role of SETs. There is evidence that teachers who are continually provided with constructive feedback from validated SET instruments will show measurable improvements in teaching skills, engagement within the classroom, and student satisfaction (Cook et al., 2022). Incorporating targeted interventions, based on the assessment of the SET, can strengthen various aspects of learning improvement, such as the use of digital tools or assessment strategy, which can boost students' learning outcomes (Khassawneh et al., 2022). The iterative feedback process can be used to build institutional capacity, create a culture of reflective practices and continuous improvement, and benefit individual instructors.

There have been several recent studies on cross-cultural and institutional differences in the use of SET. Quansah et al. (2024) highlight the importance of adapting SET instruments to local educational contexts, considering factors such as cultural expectations, teaching norms, and language differences. Likewise, the studies conducted by Heffernan et al. (2022) and AlAli & Al-Barakat (2022) highlight the importance of conducting localized validation studies to confirm the validity of the SET in various educational contexts. These adaptations are necessary to ensure reliable and interpretable findings which can be used to guide policy and practice at the institutional level.

Lastly, there is a potential path of growth in the application of AI and learning analytics in SETs in higher education research (Hooda et al., 2022). Using AI tools to analyze student feedback enables real-time assessment of classroom practices, reveal patterns in student perceptions, and offer actionable insights. These tools support traditional SETs, but they provide more granularity, timeliness, and readability in feedback, which aids in effective instructional improvement. The integration of traditional assessments with data analysis has been shown to improve the assessment of teaching effectiveness and create a comprehensive system for evidence-based decision making in education (Almulla, 2023).

The literature underscores the important function of student assessment of teaching effectiveness for evaluating faculty performance, developing faculty and enhancing student learning. Recent studies highlight the need for psychometrically sound instruments that are reliable, contextually appropriate and invariant across student groups. The use of several dimensions of teaching effectiveness such as planning, assessment, engagement, and digital competence in the development of SETs provides for the contemporary pedagogical practices. In addition, there is evidence that when well designed and well validated, they can be used to evaluate instruction but also to ensure a continuous improvement in instruction at both the faculty and institutional levels, as well as for individual faculty development. The present study seeks to contribute to the existing literature by customizing and adapting a SET instrument for the University of Sargodha Education Department, investigating cross semester and program level variations and analysing the correlation of the teaching and learning outcomes.

## **THEORETICAL FRAMEWORK**

The study is grounded on Constructivist Learning Theory which emphasizes that learning is an active process which is student-centred, where knowledge is built up through experience, reflection and interaction with the environment (Vygotsky, 1978). All variables in this research are valid and justified in relation to the concept of the effective teacher as making planning and didactic strategies, assessment and feedback, and involving students in learning, work together to reach learning outcomes. This theory proposes that instructors act as facilitators who help students synthesize new knowledge and information with what they already know, plan for meaningful learning, and facilitate critical thinking.

The didactics of planning and teaching are connected to constructivism because of well-structured, contextually relevant learning experiences. Assessment and feedback is linked to the theory's emphasis on reflective learning, with students building understanding based on formative and summative assessment. Student engagement shows the social and collaborative aspects of constructivism, presenting the significance of the interaction of peers and communication in the process of knowledge building. This study introduces the operationalization of constructivist principles by validating SET instruments for these dimensions to see how the behaviors and strategies of educators during their teaching process enable students to learn and actively construct knowledge. Therefore, constructivism offers a holistic theory model for the incorporation of teaching, assessment and student perception in higher education.

### **Conceptual Framework**

The conceptual framework is presented conceptually, showing the interconnection of the independent variables: Planning & Teaching Didactics, Assessment & Feedback, with the dependent variable: Students' Perceived Teaching Performance. Student engagement is a moderating and mediating variable between teaching strategies, assessment practices and overall teaching performance. The variables used for grouping for the purposes of examining differences between student cohorts are semester and program level. This framework validates the use of multi-dimensional SET instruments and facilitates analyses such as t-tests, ANOVA, correlation and regression, in order to gain insight into the effects of instructional implementation on student perceptions and outcomes.

### **RESEARCH DESIGN AND METHODOLOGY**

The research method for this study was quantitative with a cross-sectional research design to explore students' perceptions of the level of good teaching in Higher education. A descriptive-correlational method was used to validate the student evaluation of teaching (SET) instrument, to verify its factorial structure and to gain insight into any differences between levels and semesters in the program. Comparative analyses like t test, ANOVA, correlation and regression were used to examine relationships between teaching dimensions, student engagement, and teaching perceived performance overall. The design enabled to collect numerical data which could be statistically analysed to assess the reliability, validity and measurement invariance of the adapted SET instrument. The study took a snapshot of students' perceptions at this one point in time, ensuring that these perceptions informed faculty development and instructional improvement planning in the Education Department.

### **Population**

The population comprised all students studying in social sciences departments (Education, Sociology, Psychology, Political Science and Economics) in the University of Sargodha. The use of this inclusive population frame allowed for a connection of teaching assessment with the social sciences as a whole, and for the measurement of teaching differences between disciplines and in teaching methods. The population was important for the instrument to be validated due to its diversity of the students' experiences in social sciences. But for practical implications and for relevance to teacher training, the study was limited to the students of the Education Department, especially those taking the final semesters of studies. The evaluations of terminal semester students were viewed as those of future teachers who had taken the entire course of study and had had many contacts with faculty, and were believed to reflect the quality of instruction in teacher education classes.

### Sample and Sampling Technique

The students of Education department in the terminal semesters of the BS, MPhil and PhD of Education department were selected for the sample. Those students were chosen because they were potential teachers and had been through the entire curriculum to be able to make full assessments of teachers' performance. This study adopted purposive sampling technique, where the individuals selected were those who could provide the information that is relevant and necessary to the research goals; in this case, the terminal semester students were selected so that the information they provide is relevant and consistent. A total of 170 students participated in the final sample, representing undergraduate, postgraduate and doctoral students. This sampling method enabled the researcher to choose a particular segment of the population, yet keep the process workable. Demographic data included gender and age to describe the sample and to determine potential differences between groups.

### Adapted Instrument

The study used a Student Evaluation of Teaching (SET) adapted from Henriquez Ritchie et al. (2025) and modified for the Education Department, University of Sargodha to align with teacher education courses. The adapted instrument demonstrated high reliability with a Cronbach's alpha of 0.92 and was subjected to Confirmatory Factor Analysis (CFA) to confirm construct validity in the local context.

### DATA ANALYSIS

**Table 1: Demographic Characteristics of Respondents**

Variable	Category	f	%
Program	BS Education	115	67.6
	MPhil Education	30	17.6
	PhD Education	25	14.7
Semester	BS Education 4th	60	35.3
	BS Education 8th	55	32.4
	MPhil Education 2nd	30	17.6
	PhD Education 2nd	25	14.7
Age Range	20–22	60	35.3
	23–25	55	32.4
	24–26	30	17.6
	26–30	25	14.7
Gender	Male	80	47.1
	Female	90	52.9
Total		170	100.0

Table 1 presents the demographic characteristics of respondents. The subjects were a total of 170 students. Most of them came from the BS education (67.6%), MPhil (17.6%) and PhD (14.7%) programs. The age groups for which the majority of respondents were in the 20–22 and 23–25 ranges respectively. A slightly higher proportion of female students (52.9%) than males (47.1%).

**Table 2: Descriptive Statistics of Teaching Performance Dimensions**

Variable	N	Minimum	Maximum	M	SD
Planning and Teaching Didactics	170	1.72	4.59	3.82	0.53

Assessment and Feedback of Learning	170	1.14	5.00	3.74	0.81
-------------------------------------	-----	------	------	------	------

Table 2 shows the descriptive statistics of all the teaching performance dimensions. The students' perceptions of Planning and Teaching Didactics ( $M = 3.82$ ) and Assessment and Feedback of Learning ( $M = 3.74$ ) were positive on a 5-point Likert scale. This implies that teachers' planning, teaching, teaching strategies, assessment and feedback practices were good in general.

**Table 3: Independent Samples t-test by Gender**

Variable	Gender	N	M	SD	t	df	p
Planning and Teaching Didactics	Male	80	3.69	0.62	-3.07	168	.002
	Female	90	3.94	0.41			
Assessment and Feedback	Male	80	3.47	0.88	-4.20	144.55	< .001
	Female	90	3.97	0.65			

Table 3: Independent Samples t-test by Gender. Female students had significantly higher perceptions than male students in Planning and Teaching Didactics ( $t(168) = -3.07, p = .002$ ) and Assessment and Feedback ( $t(144.55) = -4.20, p < .001$ ). This illustrates the differences between the students' perceptions of the teaching performance of both male and female students, with females having better ratings for both dimensions.

**Table 4: Independent Samples t-test between BS 4th and BS 8th Semester Students**

Variable	Semester	N	M	SD	t	df	p
Planning and Teaching Didactics	BS 4th	60	3.85	0.46	4.34	113	< .001
	BS 8th	55	3.43	0.57			
Assessment and Feedback	BS 4th	60	3.53	0.80	0.34	113	.737
	BS 8th	55	3.48	0.92			

Table 4 shows independent samples t-test data for the comparison of BS 4th and BS 8th Semester students. The independent samples t-test data for comparison of BS 4th and BS 8th Semester students shown. There was significant difference between BS 4th and 8th semester students in Planning and Teaching Didactics:  $t(113) = 4.34, p < .001$  favors the 4th semester students. There was no difference found for Assessment and Feedback ( $t(113) = 0.34, p = .737$ ).

**Table 5: Independent Samples t-test between MPhil and PhD Students**

Variable	Group	N	M	SD	t	df	p
Planning and Teaching Didactics	MPhil 2nd	30	4.14	0.33	-1.23	35.60	.228
	PhD 2nd	25	4.21	0.10			
Assessment and Feedback	MPhil 2nd	30	4.07	0.46	-3.52	31.70	.001

Table 5: Independent Samples t-test between MPhil and PhD Students There was no significant difference between the MPhil and PhD students in terms of Planning and Teaching Didactics ( $t(35.60) = -1.23, p = .228$ ). However, the perception of assessment and feedback was significantly higher for the PhD students than for the MPhils ( $t(31.70) = -3.52, p = .001$ ), showing that advanced-level students' perception was more positive.

**Table 6: Correlation between Teaching Performance Dimensions**

Variable	1	2
1. Planning and Teaching Didactics	—	.825**
2. Assessment and Feedback of Learning	.825**	—

**Note.**  $p < .01$ .

Table 6 shows that There was a strong positive correlation between Planning and Teaching Didactics and Assessment and Feedback ( $r = .825$ ,  $p < .001$ ). High ratings for teachers by students for planning and instruction were associated with high ratings for assessment and feedback.

**Table 7: Regression Analysis Predicting Assessment and Feedback from Planning and Teaching Didactics**

Predictor	B	SE B	$\beta$	t	p
Constant	-1.03	0.25	—	-4.04	< .001
Planning and Teaching Didactics	1.25	0.07	.825	18.92	< .001

**Model Summary.**  $R = .825$ ,  $R^2 = .681$ , Adjusted  $R^2 = .679$ ,  $F(1, 168) = 357.90$ ,  $p < .001$ .

Table 7: Regression Analysis predicting Assessment and Feedback from Planning and Teaching Didactics Planning and Teaching Didactics significantly predicted Assessment and Feedback ( $\beta = .825$ ,  $p < .001$ ). The model accounted for 68.1% of the variance in Assessment and Feedback ( $R^2 = .681$ ), thus showing that the more positive the perceived teaching practices, the more positive the perceived assessment and feedback.

**Table 8: One-Way ANOVA across Semester and Program Groups**

Variable	Source	SS	df	MS	F	p
Planning and Teaching Didactics	Between Groups	15.10	3	5.03	25.21	< .001
	Within Groups	33.15	166	0.20		
	Total	48.25	169			
Assessment and Feedback	Between Groups	19.79	3	6.60	12.11	< .001
	Within Groups	90.42	166	0.54		
	Total	110.21	169			

Table 9: One-Way ANOVA across Semester and Program Group (Absolute Between Groups Value) The differences among the groups of semester/ programs were significant for Planning and Teaching Didactics ( $F(3, 166) = 25.21$ ,  $p < .001$ ) and Assessment and Feedback ( $F(3, 166) = 12.11$ ,  $p < .001$ ). This indicates that there is some disparity among students' perceptions of the teaching performance across program level and semester.

## DISCUSSION

The descriptive analysis of students' views of the performance of their teachers in the Education program indicates that the students generally have favorable attitudes to indicate the performance of their teachers. The mean scores for Planning and Teaching Didactics ( $M = 122.24$ ) and Assessment and Feedback of Learning ( $M = 52.29$ ) indicate high levels of satisfaction with the organization and assessment of

instruction. This is a positive trend that accords with recent research that indicated that if teachers focus on structured pedagogical approaches, their students feel that their educational quality is better and more satisfied with their academic lives (Chamorro-Atalaya et al., 2023). Clear instructional planning has been identified as a key aspect of improving the overall student experience in the context of modern university settings (Ritchie et al., 2025). In addition, positive attitudes towards didactics can be linked to the successful transfer of knowledge between the lesson plan and the classroom (Bazán-Ramírez et al., 2025).

Findings shows the results of the independent samples t-test to be significant for both teaching dimensions. The overall perception of Planning and Teaching Didactics was significantly higher among female students ( $M = 125.94$ ) than male students ( $M = 118.06$ ) with a significant  $t$ -value of  $-3.07$  ( $p = .002$ ). This result is corroborated by studies of the contemporary period that indicate that female students tend to give more favourable ratings on teacher assessments than male students (Bazán-Ramírez et al., 2021). Female students also scored higher ( $M = 55.62$ ), with a highly significant difference ( $p < .001$ ) between the sexes on this item of Assessment and Feedback. A number of recent systematic reviews have highlighted that gender of students is a continuing issue in the perception and rating of teaching quality in higher education (Stoesz et al., 2023). These differences have been linked to different expectations for the communicative support and feedback from faculty by some researchers (Tangalakis et al., 2022). The lack of improvement in these gaps underscores the need to consider demographic factors when relying on student evaluations as a sole indicator of teaching effectiveness (Valencia et al., 2025).

Findings show that the teaching perception of students' academic stages reveals that it is a determining factor. For example, the 4th semester BS students' mean of planning and didactics ( $M = 123.22$ ) was significantly higher than that of the 8th semester BS students' ( $M = 109.91$ ), giving a  $t$  value of  $4.34$  ( $p < .001$ ). This drop in satisfaction ratings as students approach graduation may be attributed to more critical evaluative standards as students become more mature and develop knowledge and expertise in their fields (Dobson et al., 2025). Interestingly, Table 5 reveals that students who are pursuing PhDs were significantly more satisfied with Assessment and Feedback ( $M = 61.28$ ) than those who are pursuing MPhil ( $M = 57.03$ ). The higher rating for doctorate students may be due to the more individual and detailed feedback that occurs with advanced research supervision (Harris et al., 2022). Table 8 also shows results of the one-way ANOVA to further validate the significant difference in perceptions of didactics ( $F = 25.21$ ,  $p < .001$ ) and assessment ( $F = 12.11$ ,  $p < .001$ ) between the programs and semester levels (Ritchie et al., 2025).

Findings show correlation and regression analyses among instructional planning and assessment satisfaction and suggest a strong relationship between instructional planning and assessment satisfaction. The Pearson correlation coefficient ( $r = .825$ ,  $p < .001$ ) suggests high interrelatedness of these two dimensions in the minds of students. The findings of recent literature have disclosed that didactic strategies have been found to be one of the most important predictors of students' views on fairness and quality in the evaluation process (Qi & Huang, 2024). In particular, Planning and Teaching Didactics accounted for 68.1% ( $R^2 = .681$ ) of the variance in Assessment and Feedback. This is aligned with recent regression models in educational research that suggest effective teaching strategies are key factors in effective feedback cycles (Morales-Romero et al., 2021). If teachers are seen to be organized and well-prepared, students will be more likely to see the assessment and feedback as meaningful and constructive (Rigopoulos, 2022). Finally, teacher-student interaction via feedback is a key factor in the student-centered education of 2020 plus (Tong et al., 2023).

Finally, the results suggest that there are differences in the perception of teaching quality according to the characteristics of students and their academic level. There are clear gender differences, and differences between undergraduate and postgraduate perceptions, which means that a 'one-size-fits-all' approach may

not be appropriate for a diverse student population. In addition, the high predictive power of planning on feedback emphasizes the need for a well-designed curriculum. Teachers need to pay more attention to a didactic clarity and to create a more positive assessment culture. The insights are useful for the University administrators who want to enhance the academic quality of the University by targeting specific needs and evaluative patterns of student groups.

## CONCLUSION

Finally, this study confirms the generally favourable attitude of education students to teaching performance, with a specific focus on planning, didactic and assessment skills. The results show that teachers are considered to be well-prepared and effective in teaching. The high gender disparity however does show that there is a difference in the assessment of teaching performance between males and females, perhaps in assessment expectations or patterns within the classroom.

Moreover, the findings in this study reveal that academic trajectory affects student perceptions. This finding is particularly true for younger students at the undergraduate level, who are more satisfied with instructional planning, than for doctoral students, who report higher levels of satisfaction with feedback and assessment mechanisms than do MPhil students. The most important observation is that instructional planning and assessment are closely linked; teachers' planning and teaching of clear lessons have a direct influence on the quality and fairness of the students' feedback. This is a strong correlation, which highlights the fact that high-quality teaching is a process of integration, and strategic planning is the basis for effective evaluation. In conclusion, the results of this study have important implications for university departments to adapt their teaching methods so that teaching quality is uniform throughout each department while meeting the needs of a diverse student body.

## REFERENCES

- AlAli, R., & Al-Barakat, A. (2022). Using SEM to assess creative teaching perceptions in higher education. *Education Sciences, 12*(10), 690. <https://doi.org/10.3390/educsci12100690>
- Almulla, M. A. (2023). Constructivism learning theory: A paradigm for students' critical thinking, creativity, and problem solving to affect academic performance in higher education. *Cogent Education, 10*(1), 2172929. <https://doi.org/10.1080/2331186X.2023.2172929>
- André, S., Maulana, R., Helms-Lorenz, M., & Telli, S. (2023). Measuring teaching behavior in higher education: Multi-group confirmatory factor analysis approach to measurement invariance. *Frontiers in Psychology, 14*, 1123456. <https://doi.org/10.3389/fpsyg.2023.1123456>
- Bazán-Ramírez, A., Henostroza-Mota, C., Peña-Correa, G. B., Capa-Luque, W., Mayorga-Falcón, L. E., Begazo-Jaramillo, A. E., & Quicaña-Amáu, F. P. (2025). Professors' didactic performance and students' didactic performance in psychology classes at a Peruvian university. *Frontiers in Education, 10*. <https://doi.org/10.3389/feduc.2025.1463493>
- Bazán-Ramírez, A., Pérez-Morán, J. C., & Bernal-Baldenebro, B. (2021). Criteria for Teaching Performance in Psychology: Invariance According to Age, Sex, and Academic Stage of Peruvian Students. *Frontiers in Psychology, 12*. <https://doi.org/10.3389/fpsyg.2021.764081>
- Berk, R. A., & Theall, M. (2023). *Thirteen strategies to measure college teaching: A consumer's guide to rating scale construction, assessment, and decision-making for faculty, administrators, and policymakers*. Taylor & Francis. <https://doi.org/10.4324/9781003448181>

- Chamorro-Atalaya, O., Pacheco, A. J. G., Silvestre, M. Q., Herrera, V. H. D., Suarez-Bazalar, R., Vargas-Díaz, A., Villanueva-Acosta, V., & Chinchay, L. R. M. (2023). Professional and personal traits of the teacher and the relationship with didactic strategies. *International Journal of Evaluation and Research in Education (IJERE)*, 13(1), 122–122. <https://doi.org/10.11591/ijere.v13i1.26957>
- Constantinou, C., & Wijnen-Meijer, M. (2022). Student evaluations of teaching and the development of a comprehensive measure of teaching effectiveness for medical schools. *BMC Medical Education*, 22, 3148. <https://doi.org/10.1186/s12909-022-03148-6>
- Cook, C., Jones, J., & Al-Twal, A. (2022). Validity and fairness of utilizing student evaluation of teaching (SET) as a primary performance measure. *Journal of Further and Higher Education*. <https://doi.org/10.1080/0309877X.2021.1895093>
- Daumiller, M., Janke, S., Hein, J., & Rinas, R. (2022). Teaching quality in higher education: Measurement and validity considerations. *Assessment*, 29(3), 321–342. <https://doi.org/10.1027/1015-5759/a000700>
- Dobson, B., Ward, S., Corden, S., & Tinnion, D. J. (2025). Comparative analysis of National Student Survey responses: Exploring the perspectives and feedback of undergraduate sport students. *Journal of Hospitality Leisure Sport & Tourism Education*, 36, 100550–100550. <https://doi.org/10.1016/j.jhlste.2025.100550>
- Fernández-García, C. M., Maulana, R., & Helms-Lorenz, M. (2023). Cross-cultural measurement invariance of student evaluation of teaching in six countries. *Frontiers in Psychology*, 14, 1156792. <https://doi.org/10.3389/fpsyg.2023.1156792>
- Harris, R., Blundell-Birtill, P., & Pownall, M. (2022). Development and evaluation of two interventions to improve students' reflection on feedback. *Assessment & Evaluation in Higher Education*, 48(5), 672–685. <https://doi.org/10.1080/02602938.2022.2107999>
- Heffernan, T., Smith, R., & Johnson, P. (2022). Sexism, racism, prejudice, and bias: A literature review and synthesis of research surrounding student evaluations of courses and teaching. *Assessment & Evaluation in Higher Education*, 47(8), 1056–1078. <https://doi.org/10.1080/02602938.2021.1888075>
- Hooda, M., Rana, C., Dahiya, O., & Rizwan, A. (2022). Artificial intelligence for assessment and feedback to enhance student success in higher education. *Mathematical Problems in Engineering*, 2022, 5215722. <https://doi.org/10.1155/2022/5215722>
- Khassawneh, O., Mohammad, T., & Ben-Abdallah, R. (2022). The relationship between emotional intelligence and educators' performance in higher education sector. *Behavioral Sciences*, 12(12), 511. <https://doi.org/10.3390/bs12120511>
- Khlaif, Z. N., Ayyoub, A., Hamamra, B., & Bensalem, E. (2024). University teachers' views on the adoption and integration of generative AI tools for student assessment in higher education. *Education Sciences*, 14(10), 1090. <https://doi.org/10.3390/educsci14101090>
- Lakeman, R., Coutts, R., & Hutchinson, M. (2022). Appearance, insults, allegations, blame, and threats in student evaluations. *Assessment in Higher Education*. <https://doi.org/10.1080/02602938.2021.2012643>

- Maki, P. L., & Kuh, G. D. (2023). *Real-time student assessment: Meeting the imperative for improved time to degree, closing the opportunity gap, and assuring student competencies for 21st century*. Taylor & Francis. <https://doi.org/10.4324/9781003446675>
- Marsh, H. W., & Hocevar, D. (2023). Factorial validity and invariance of student evaluations of teaching in higher education. *American Educational Research Journal*, 60(1), 45–72. <https://doi.org/10.3102/0002831223112345>
- Okoye, K., & Arrona-Palacios, A. (2022). Towards teaching analytics: Evaluation of student feedback through text mining. *Education and Information Technologies*, 27, 5671–5689. <https://doi.org/10.1007/S10639-021-10751-5>
- Qi, S., & Huang, T. (2024). Using Data Analysis to Improve the Teaching Quality of Civics in College English Courses. *Applied Mathematics and Nonlinear Sciences*, 9(1). <https://doi.org/10.2478/amns-2024-1377>
- Quansah, F., Cobbinah, A., & Asamoah-Gyimah, K. (2024). Validity of student evaluation of teaching in higher education: A systematic review. *Frontiers in Education*, 9, 1329734. <https://doi.org/10.3389/feduc.2024.1329734>
- Rigopoulos, G. (2022). Assessment and Feedback as Predictors for Student Satisfaction in UK Higher Education. *International Journal of Modern Education and Computer Science*, 14(5), 1–9. <https://doi.org/10.5815/ijmeecs.2022.05.01>
- Ritchie, P. S. H., Pérez-Morán, J. C., García, C. J. del C., & Cervantes, B. B. (2025). Evaluation of teaching performance by university students of social and administrative sciences: a comparative study. *Frontiers in Education*, 10. <https://doi.org/10.3389/feduc.2025.1629540>
- Stoesz, B. M., Jaeger, A. E. D., Quesnel, M., Bhojwani, D., & Los, R. (2023). Bias in Student Ratings of Instruction: A Systematic Review of Research from 2012 to 2021 [Review of *Bias in Student Ratings of Instruction: A Systematic Review of Research from 2012 to 2021*]. *Canadian Journal of Educational Administration and Policy*, 201, 39–62. <https://doi.org/10.7202/1095482ar>
- Tong, Y., Zhang, L., & Zhu, Y. (2023). Research on the Learning Status of Students in High-Level Application Oriented Universities Based on Student-Centered Theory. *Open Journal of Social Sciences*, 11(1), 11–22. <https://doi.org/10.4236/jss.2023.111002>
- Valencia, E., Flores, A., & Navarro-Ibáñez, M. (2025). What Have We Learned About the Instructor's Gender Effect on Student Evaluation of Teaching From Experimental and Quasi-Experimental Research? *Review of Educational Research*. <https://doi.org/10.3102/00346543251379479>
- Xia, Q., Weng, X., Ouyang, F., Lin, T. J., & Chiu, T. K. F. (2024). AI-enhanced assessment in higher education. *International Journal of Educational Technology in Higher Education*, 21, 468. <https://doi.org/10.1186/s41239-024-00468-z>