

Evaluating the Teaching Performance in the English Language Proficiency Program at Al-Ahgaff University in the Light of Quality Assurance Standards from Students' Perspectives

Mohammed Salim Karamah Al-Bouri

PhD Researcher at the Department of English, Faculty of Education, Sana'a University, Yemen

m.s.k.alboori@gmail.com

Fazee Khalid Alezi Mohammed Almuslimi

Professor of Applied Linguistics (Curricula and Methods of Teaching English), the Head of English Department, Faculty of Education, Sana'a University, Yemen

aboehab1977@gmail.com

Abdullah Rajab Alfalagg

Assist. Prof. of TESOL, Hadhramout University, Yemen

a.alfalagg@hu.edu.ye

DOI: <http://doi.org/10.36892/ijlls.v7i6.2412>

APA Citation: Al-Bouri, M. S. K., Almuslimi, K. A. M., & Alfalagg, A. R. A. (2025). Evaluating the Teaching Performance in the English Language Proficiency Program at Al-Ahgaff University in the Light of Quality Assurance Standards from Students' Perspectives. *International Journal of Language and Literary Studies*. 7(6).266-287. <http://doi.org/10.36892/ijlls.v7i6.2412>

Received:

16/09/2025

Accepted:

29/10/2025

Keywords:

Evaluation,
Teaching
Performance,
Perspectives,
Quality
Assurance
Standards,
English
Language
Proficiency
Programs,
Tertiary level

Abstract

English Language Proficiency Programs (ELPPs) have proliferated in Yemeni tertiary institutions due to the Ministry of Higher Education's requirement that all students intending to enroll in graduate programs in Yemen should obtain certification in the English language. The present study aims to evaluate the teaching performance in the ELPP at Al-Ahgaff University in light of quality assurance standards. It investigates students' perceptions of implementing the indicators of teaching performance based on quality assurance standards in the ELPP. This study employed a descriptive quantitative method, utilizing a questionnaire distributed to the participants to assess the implementation of quality assurance standards in teaching performance. To achieve the objectives of the study, 162 students responded to a closed-ended questionnaire in which the collected data were quantitatively analysed through the SPSS program. Descriptive and inferential statistical measures were computed to analyze the respondents' responses. The findings showed that (10) of the teaching performance indicators were as highly implemented while the other six indicators were moderately implemented by the teachers in the ELPP. The results also revealed significant differences in the students' perceptions of implementing the indicators attributed to gender, in favor of female students. However, no statistically significant differences appeared due to the participants' level of study. Such significant implications were categorized based on resource demand and potential impact for developing English language proficiency programs.

1. INTRODUCTION

Although English is taught as a foreign language in Yemen, it gains high significance at the tertiary level. English is used in instruction for specific purposes (ESP) and for academic purposes (EAP) to teach scientific content (Alfalagg, 2018; Alotumi, 2015; Ba-udhan, 2014).

Quality Assurance (QA) in education is essential in ensuring the effectiveness of teaching and learning processes, particularly in language programs where performance and pedagogical competence are essential (Harvey & Green, 1993). The performance of English language teachers plays a pivotal role in determining the success of language programs, as they are directly responsible for delivering curriculum content, fostering students' engagement, and achieving learning outcomes (Richards, 2017). Evaluating teachers' performance in light of quality assurance standards provides a structured framework for assessing teaching effectiveness, identifying areas for improvement, and enhancing overall program quality (European Association for Quality Assurance in Higher Education [ENQA], 2015).

Quality assurance standards in education typically encompass key dimensions such as curriculum design, teaching methodologies, assessment practices, professional development, and learner feedback (British Council, 2018). These standards evaluate whether English language teachers meet the required competencies and contribute to program excellence. Research indicates that systematic teachers' evaluation enhances instructional strategies, improves students' outcomes, and fosters accountability in educational institutions (Darling-Hammond, 2017).

According to QA standards, students' perspectives are essentially incorporated in the teaching process of educational programs, competent teachers, textbooks etc. Students' feedback offers direct insights into teaching effectiveness, curriculum relevance, learning experiences, and institutional support, all of which are essential for maintaining high academic standards (Berk, 2018). According to Marsh (2007), students' evaluations help in identifying strengths and weaknesses in the teaching process, course content, and assessment practices. Different researches show that students' feedback leads to pedagogical improvements, fostering better learning outcomes. When students feel their voices heard, their engagement and satisfaction increase. Surveys reveal that institutions acting on students' feedback see higher retention rates (Ameyaw & Khumalo, 2024)

Richards (2021) suggests that English language proficiency programs necessitate systematic evaluation to meet learners' needs while maintaining high educational standards effectively. Regular assessment of these programs is crucial for verifying learning outcomes, improving instructional quality, and demonstrating institutional accountability (Richards, 2017). As English continues to serve as the global lingua franca (Crystal, 2012), rigorous assessment of these programs becomes increasingly vital for educational institutions worldwide.

1.1.Statement of the Problem

Yemen's ongoing civil war has caused catastrophic disruptions across social, economic, and political spheres, leading to severe declines in educational attainment (UNICEF, 2023). Hence, the significance of the present study lies in evaluating teachers' performance and enhancing it to meet the national and international standards of institutional quality. Furthermore, English language proficiency programs at higher education institutions have gained popularity in Yemen, as several studies recommend establishing such programs to enhance the language competence of incoming college students (Alfalagg & Ba-Udhan, 2023). Language proficiency programs have long been in place to help overseas students acquire the academic credentials needed for admission, but often lack the necessary language skills (Crosthwaite, 2016; Keefe & Shi, 2017; Storch & Tapper, 2009).

Students' perspectives play a crucial role in assessing teachers' performance in English language proficiency programs, as students are the immediate beneficiaries of instructional quality (Marsh, 2007). However, many institutions lack structured mechanisms to incorporate students' feedback into QA frameworks, leading to potential gaps in evaluating teaching effectiveness (Spooren et al., 2013). Existing research on teachers' evaluation tends to focus on institutional and peers' reviews rather than student-centered assessments aligned with QA benchmarks (Penny & Coe, 2004). Carless (2019) states that there is a limited exploration of how students' feedback can be systematically integrated into QA frameworks to enhance English teachers' performance evaluations while maintaining objectivity and fairness.

English language policy at Al-Ahgaff University emphasizes the use of the English language in teaching across different faculties. Therefore, the need for an intensive English proficiency program has become a matter of great urgency. On this basis, the ELPP program was established in 2016 at Al-Ahgaff University. It has been observed that the deans of different faculties, instructors, and students express dissatisfaction with the ELPP. Although Alaidaros et al. (2021) have evaluated academic programs at Al-Ahqaf University, no research has investigated the teaching performance and the quality of the ELPP outcomes since its the establishment. Previous research has examined the development of quality assurance systems and accreditation mechanisms in Yemeni higher education, focusing on the implementation of quality assurance standards (Alguhali et al., 2022; Al-Muslimi, 2011; Al-Saba et al., 2010; Ateeq, 2023; Jaber, 2021). However, no study to date has investigated teaching performance indicators in English language proficiency programs, such as metrics for classroom instruction quality, faculty feedback mechanisms, or learner-centered teaching evaluations. Such the above mentioned gap presents one of the motives for conducting the present evaluative study.

1.2. Significance of the Study

The present study aims to evaluate the performance of English language teachers within the English language proficiency program (ELPP) approved by Al-Ahgaff University in light of QA standards by applying established standards. It would be of paramount importance to stakeholders at Al-Ahgaff University, particularly decision-makers and the office of quality assurance and accreditation. By examining teaching practices against quality assurance standards, the study aims to provide insights into the program's strengths, weaknesses, and potential areas for improvement. The results might help evaluate other programs at Al Ahgaff University based on quality assurance standards, such as the Arabic language learning program for non-native speakers. The findings could contribute to a broader discourse on teachers' evaluation and quality assurance in English language proficiency programs. Its findings might apply to public universities in Yemen that offer English language proficiency programs.

1.3. Delimitations of the study

The current study was limited to evaluate the teaching performance of the teachers in the English Language Proficiency Program (ELPP) in light of quality assurance standards. It focused on the students' perceptions of teachers' teaching performance. The students were enrolled in the Faculty of Administration and Economics and the Faculty of Computer Sciences

where they had to simultaneously take the ELPP as a prerequisite course. Moreover, the study was conducted at Al-Ahgaff University in Mukalla during the 2024–2025 academic year

1.4. Questions of the Study

- 1- To what extent are the indicators of teaching performance implemented in the ELPP based on the quality assurance standards from the students' perceptions?
- 2- Is there a statistically significant difference in the students' perceptions of implanting the indicators of the teaching performance due to the variables of gender and the level of study?

2. THEORITICAL BACKGROUND

2.1. Quality and Quality Assurance in Higher Education

Most quality assurance procedures were originally developed for the industrial sector and subsequently adapted for the service sector, which includes education. The history of quality assurance in higher education has followed a similar trajectory to that of the industry (Koslowski, 2006). Voreijensijn (1995) outlined that quality assurance in higher education has been present since the inception of universities, with its origins dating back to the medieval era. During the 1990s, the notion of quality was prominent within Higher Education Institutions (HEIs). Over time, there was a growing awareness of its essential nature, culminating in the establishment of various departments, units, and networks equipped with diverse methodologies and instruments for Quality Assurance (QA) (Usmani, Khatoon, Shammot, & Zamil, 2012).

Various interpretations and definitions of quality assurance have been introduced in the literature. To begin with, Vlasceanu, Grünberg and Pârlea (2005) defined quality assurance as an ongoing process of evaluation that includes monitoring, assessing, maintaining, and improving the quality of higher education programs, institutions, or systems. Similarly, Petersen (1999) referred to quality assurance in higher education as all the purposeful and methodical actions required to instill trust by which international instruction, scholarship, and qualification standards are being maintained and improved. The word quality is derived from the Latin word *qualis*, which means "what kind of" (Ali & Shastri, 2010, p. 9). Quality is a concept that is difficult to define and is also a term that carries subjective value. Similarly, Watty (2006) noted that the term "quality" has become a subject of debate and that there is considerable confusion regarding its precise definition. It is commonly linked to what is considered good and valuable (Green, 1994). Most analysts and policymakers in higher education referred to quality as fitness for purpose (Redder, 2010).

Harvey and Green (1993) emphasized the complex nature of defining quality within the higher education domain, due to the varying interpretations that different stakeholders associate with it (Rasool, 2010). Harvey and Green (1993) emphasized that quality is evaluated based on the degree to which the product or service aligns with its intended purpose. If the product fails to adhere to its intended purpose, its state of perfection becomes irrelevant. According to Vlasceanu et al. (2007), quality as fitness of purpose concentrates on the stated goals and mission of a program or institution without evaluating the processes' suitability in relation to

any outside goals or expectations. They added that the appropriateness of purpose assesses the adequacy of quality-related activities of the institution (Vlasceanu et al., 2007). In addition to adding a new dimension to the definition of quality in higher education, Campbell and Rozsnyai (2002) have replaced the term "fitness for purpose" with "quality as threshold." This means that academicians are the best people to define their own course of action because they possess the most knowledge about quality in higher education. All the aforementioned definitions of quality show that quality is a philosophical concept. It varies and, to some extent, reflects different perspectives of the individual and society.

2.2. Language Program Evaluation

In educational settings, the examination of program evaluation has emerged as a significant area of scholarly inquiry, wherein Language Program Evaluation (LPE) has assumed a crucial role (Uzun, 2016). Within educational environments, the exploration of program evaluation has emerged as a prominent domain of academic investigation, in which Language Program Evaluation (LPE) has taken on a pivotal role (Uzun, 2016). In recent years, the evaluation of programs has garnered the attention of scholars worldwide and has been recognized as an essential area of inquiry within educational research (Harris, 2009; Norris, 2009; Owen, 2007; Patton, 2002; Peacock, 2009; Uzun, 2016). As Brown (1995) indicated, one of the three essential elements constituting the ongoing curriculum evaluation process is the continuous assessment of programs. Its roles may encompass a range of activities, from formative internal improvements to comprehensive summative assessments (Harris, 2009) that aim to enhance the program, ensure educational effectiveness, and promote the sustainability of the language instruction profession (Norris, 2009). The Language Program Evaluation (LPE) benefits from specific fields within applied linguistics, including second language acquisition, classroom-focused research, and language assessment. Additionally, it draws on perspectives from the social and political foundations of language usage and philosophical investigations into science (Lynch, 1996).

Furthermore, LPE has evolved significantly beyond the simplistic confines of accountability and outcome-focused considerations. It now includes a more comprehensive framework that comprises 'framework, heuristic, and methodologies for interpreting language teaching and learning (and other language-related activities) in situ, as well as for addressing 'what works' inquiries and guiding practical initiatives for enhancement' (Norris, 2016). Additionally, within the context of scholarly analysis, there has been a considerable focus on evaluating educational or language programs. This field of research occupies a fundamental position in the progress of curriculum formulation and the safeguarding of educational integrity (Brown, 1989; Norris, 2016).

Teaching performance evaluation in language programs is a critical component of quality assurance, ensuring instructors meet pedagogical, linguistic, and cultural competency standards (TESOL International Association, 2019). In English language proficiency programs, evaluations must also account for instructors' ability to foster communicative competence and adapt to diverse learner needs (Richards & Farrell, 2005). To achieve the goals of effective teaching in EFL, teachers require a combination of strong teaching skills, in-depth knowledge of their subject, and high proficiency in the language itself (Bi, 2012). Effective evaluations

typically incorporate multiple assessment methods, including classroom observations, self-assessments, and student feedback to measure teaching effectiveness holistically (Darling-Hammond, 2017).

2.3. Students' Evaluation of Teaching Performance

Teaching performance refers to observing and evaluating a teacher's skills in actually carrying out an activity or producing a product within the classroom (Airsian, 2000). Teachers' performance plays a key role in the quality of education by influencing how effectively the teaching and learning process is managed (Aimah & Purwanto, 2019). Research shows that effective teaching relies on teachers having specific knowledge, skills, and experience (Santiago, Roseveare, van Amelsvoort, Manzi, & Matthews, 2009; Syahrudin et al., 2013). In the present study, teaching performance is defined as the level of quality and proficiency demonstrated by the teacher in implementing instructional methods and strategies to achieve the prescribed educational objectives. This includes clarity in presenting content, effective lesson organization, efficient time and classroom management, active student interaction, the use of strategies tailored to learners' needs, and the application of assessment methods to support learning. Teaching performance is measured through observable and assessable indicators, such as student evaluation questionnaires, peer observations, and assessment tools based on quality standards adopted in higher education.

Students' evaluations refer to their feedback and ratings of instruction, and they have been a widespread practice on college campuses for decades. Despite some issues, students' evaluations are generally accepted by both researchers and practitioners (Marsh, 1987). They provide direct measures of consumer satisfaction with teaching. Previous research studies attest that students' evaluation of teaching performance is useful and considerably accurate because students interact with teaching in a first-hand experience more than any other observer. Therefore, they are in an unrivalled position to share their observation (Ramsden, 1991). Students' evaluations of teaching provide reliable data for assessing instructional quality and curriculum effectiveness (Marsh, 1987). When systematically collected and analyzed, students' feedback helps institutions identify strengths and areas for improvement in pedagogical approaches, learning materials, and classroom dynamics (Carless, 2019). Interestingly, aligning students' evaluation with international QA standards helps ensure assessment consistency and fairness (Council of Europe, 2020).

The validity of students' evaluations of teaching performance is supported by the existence of positive relationships between student evaluations and students' achievement. Well-designed ratings instruments primarily reflect the instructor's performance. They are considered reliable for measuring effective teaching and are considerably unaffected by potential biases (Marsh, 1993; Ramsden, 1991). Teachers consider students' feedback as a valuable indicator of the quality of their teaching performance and the curriculum (Surujlal, 2014). Despite its significant contribution to program development, students' evaluation of teaching performance is paradoxical.

Integrating students' evaluation of teaching performance into QA frameworks often faces challenges. Defects in students' evaluation include variability in students' perspectives, cultural

differences in feedback styles, and the potential for bias (Carless, 2019; Spooren et al., 2017). On this basis, its reliability is low to depend entirely on for assessing teacher performance (Ramsden, 1991), especially when evaluations are tied to high-stakes teacher assessments. A key challenge in students evaluating language teachers lies in balancing objectivity (Hattie, 2017). For instance, Arnon and Reichel (2007) and Ida (2017) suggest that students often rate teachers based more on personality than on teaching skills. Students' potential bias could be minimized by controlling the methods of administration. Students might be asked to comment on aspects of teaching in which they are qualified (Ramsden, 1991). Many programs combine student evaluations with peer reviews, classroom observations, and learning outcome analyses to enhance reliability, ensuring a more balanced and objective assessment of teaching quality (Darling-Hammond, 2012; Garrett & Steinberg, 2015; Seidel & Shavelson, 2007).

2.4. The English Language Proficiency Program (ELPP) at Al-Ahgaff University

In line with Al-Ahgaff University's mission, ELPP was established in 2016, and is catered to non-English major students. It is designed for students enrolled in the Faculty of Administration and Economics, as well as the Faculty of Computer Sciences. The ELPP curriculum is a hybrid of general English and specialized materials tailored to students' fields of study. The ELPP spans four semesters over a two-year period. It comprises a total of 300 contact hours over 52 weeks. Before the beginning of the first semester, the admitted students to the program undergo a placement test to assess their proficiency level in the English language. Accordingly, they are placed in appropriate levels within the ELPP. Students may be exempt from the program if they achieve a minimum score of 400 on the TOEFL or an equivalent IELTS score. The ELPP aims to equip students with the English language skills necessary to learn content courses in their mainstream specializations at the affiliated colleges of Al-Ahgaff University (English Language Unit, 2016). The primary goal of the ELPP is to equip university students with the necessary competencies in the English language required for their academic studies based on their respective majors of accounting and information technology. The program emphasizes key competencies, including academic reading, academic writing, listening to lectures on academic content, and speaking in an academic context.

2.5. Previous Studies

Several studies have examined students' perceptions of teachers' performance in English proficiency programs, focusing on how these programs align with quality assurance (QA) standards. Key factors include teaching effectiveness, language competency, assessment methods, and institutional support.

To begin with, Abdulsalam and Al-Qataani (2024) evaluated the effectiveness of faculty performance at Ajdabiya University, Faculty of Arts using quality standards that encompassed course planning, delivery, personal attributes, student interaction, and assessment. Results revealed a moderate overall performance, with student interaction skills rated as the highest. Statistically significant variations emerged based on gender (favoring female faculty), discipline (geography ranked highest), and academic level (third-semester students provided the most positive evaluations). These findings underscore discipline-specific and experiential

factors in teaching effectiveness assessments, suggesting the need for tailored faculty development programs to address performance gaps.

Al-Rsa'i et al. (2024) investigated the quality of teaching performance among 347 teachers using a 40-item measure across three domains: planning, implementation, and assessment. The study analyzed the impact of gender, experience, teaching stage, and training on performance. The results showed that teachers' performance was generally average, but their performance in the first three grades was the weakest. At the same time, it was also shown that the female teachers' performance was superior to that of male teachers. The performance of specialized teachers in the scientific field was better than that of their counterparts in the humanitarian field. The results showed the positive impact of professional experience and training on the quality of teaching performance. The study also demonstrated the importance of the first five years in the teachers' work. Therefore, focusing on training and qualification was recommended, as well as providing appropriate training for pre-service teachers.

Al-Maraziq and Bani (2022) evaluated the quality of academic programs within Jerash University's College of Educational Sciences, as perceived by its undergraduate and graduate students. The findings indicated a generally high student assessment of program quality across the curriculum and syllabus domain, as well as the faculty members' domain. Faculty members received the highest ratings, followed by curriculum and syllabi. While student perceptions did not vary by gender, MA students rated program quality significantly higher than BA students.

Johari (2021) evaluated university teaching performance from students' perspectives in light of educational quality assurance standards. The study results showed that the level of university teaching performance, as rated according to educational quality assurance standards, was evaluated as average by social sciences students. No statistically significant differences in social sciences students' evaluations of university teaching performance were attributed to gender.

Faraj and Al-Anqari (2020) investigated the extent to which quality standards are implemented in the teaching performance of faculty members from the perspective of college students at the College of Education. The results showed that the availability of quality standards in teaching performance was moderate, with statistically significant differences in favor of applied sciences.

Al-Mekhlafi (2020) assessed the implementation of quality standards in graduate programs at the College of Education, Imam Abdulrahman Bin Faisal University (IAU), based on the National Center for Academic Accreditation and Evaluation (NCAAA) standards from students' perspectives. The findings revealed an overall moderate level of implementation, with higher compliance in areas related to teaching staff and quality learning. However, lower performance was observed in standards concerning scientific research and projects, students, learning resources, facilities and equipment, curricula, mission and goals, as well as graduate qualities and learning outcomes. Additionally, a statistically significant gender difference was found in perceptions of teaching staff quality, favoring female students.

Darling-Hammond (2017) examined how quality assurance standards for English language programs should encompass both teaching competencies and student outcomes. The study

proposed a framework linking teacher performance metrics to program accreditation standards. In addition, the study highlighted the need to link teachers' performance evaluations with students' grades to ensure quality. Marsh (2007) at Australian universities analyzed 15,000 student evaluations of English language teachers, identifying key QA indicators. The study showed that instructional clarity, engagement strategies, and feedback effectiveness are the most important quality assurance measures from the students' perspectives.

Al-Subaie (2010) investigated the extent to which female faculty members employed science teaching skills aligned with comprehensive quality standards, as perceived by their students. Findings revealed that students perceived faculty members' implementation of quality-standard-aligned science teaching practices as generally inadequate. However, the results showed disciplinary variations, with faculty in the biology department demonstrating significantly stronger performance in the teaching competencies compared to their counterparts in other science disciplines.

Canaan (2005) conducted a study to enhance faculty members' performance based on total quality standards. The study aims at determining the extent to which current teaching performance among faculty members aligns with these comprehensive quality standards. It identified specific indicators that were fully achieved, including selecting appropriate assessment methods, involvement in curriculum and course development, and supervision of Master's and PhD theses.

Zniati (2005), as cited in Abu-Al-Sha'r and Al-Harabsheh (2013) examined the fundamental principles for evaluating faculty at the University of Damascus, using quality standards and accreditation systems as benchmarks. The key findings included the faculty workload distribution, sabbatical opportunities, and the effective integration of modern instructional technologies. The study further demonstrated the institution's strong commitment to excellence, as evidenced by its significant progress in enhancing the quality of its performance. These advancements positioned the college favorably among global dental institutions in terms of compliance with comprehensive quality and accreditation standards.

Alsirr (2004) conducted a study to assess teaching competency among faculty members at Al-Aqsa University in Gaza. Findings revealed that overall teaching skills met established quality benchmarks, particularly in instructional planning (82%) and classroom communication (80%). However, assessment competencies fell below quality standards. The study identified academic qualification as a significant factor influencing teaching and evaluation quality, while neither professional experience nor faculty affiliation demonstrated measurable effects on instructional performance.

Al-Hakmi (2003) aimed to establish a benchmark for the essential professional competencies of university professors based on students' viewpoints and to identify the most valued competencies at Umm Al-Qura University. The findings highlighted that students prioritized the following key areas: personal traits, lecture preparation and delivery, interpersonal skills, extracurricular engagement and assessment, subject-matter expertise and professional development, and motivational and reinforcement techniques.

Alkyumi (2002) investigated the application of specific quality management principles to teacher training colleges in Oman. Results showed that administrators and faculty members had similar perceptions regarding the applicability of certain total quality management concepts. Additionally, most areas demonstrated a high level of applicability, except for the use of scientific methods, which was rated at a moderate level. The study examined key aspects, including institutional mission, decision-making processes, delegation of authority, autonomy, scientific methodology, teamwork, professional development, and ongoing evaluation.

This review of previous studies indicates that there is interest in applying total quality assurance in various educational institutions. Moreover, there are orientations toward the serious application of comprehensive quality standards in university education, including inputs, processes, and outputs, utilizing regulatory total quality management frameworks. Most of the aforementioned studies focus on the reality of applying total quality standards in educational institutions globally.

3. RESEARCH METHOD AND DESIGN

The study adopted descriptive quantitative research method. As per data collection, the study design was cross-sectional (Creswell & Creswell, 2018) by contacting the participants at one time and requested that they respond to a questionnaire. This design was chosen for its efficiency and practicality in capturing a snapshot of participants' perceptions at a single point in time, aligning well with the study purpose of evaluating the implementation of teaching performance indicators. Cross-sectional design provides a time-effective means to analyze and understand the current state of teaching performance in the ELPP from the students' perspective. The quantitative data were collected through a close-ended questionnaire at one time.

3.1. Population and Sample of the Study

All students enrolled in the ELPP at Al-Ahgaff University constituted the population of this study. They were 300 students majoring in computer sciences and administration. The ELPP is a perquisite course that qualify students to be promoted to level 3 at Al-Ahgaff University. The sample consisted of (162) students randomly selected based on the attendance sheet of the ELPP. One hundred twenty-two of the participants (75.3%) were males, and 40 of them (24.7%) were females. Eighty of them (49.4%) were studying at the first level in college, and 82 of them (59.6%) at the second level.

Table 1. The Participants' Demographic Variables

Variables	Category	Frequency	Percentages
Gender	Male	122	75.3
	Female	40	24.7
Level of study at college	First level	80	49.4
	Second level	82	50.6

3.2. Data Collection Instrument

For the purpose of achieving the aims of the study, a questionnaire with three parts was used to explore the extent to which quality assurance standards were applied in the teaching performance of the ELPP from the students' perspective. The introductory part of the questionnaire contained demographic information such as gender, age, level of study, and department. The second part contained (22) close-ended indicators based on the previous literature (Al-Maraziq & Bani, 2020; Al-Mahrooqi et al., 2015; Al-Mekhlafi, 2020; Ataman & Adıgüzel, 2024; Faraj & Al-Anqari, 2020; Quach & Nguyen, 2025; Ramsden, 1991). The indicators highlight the students' perceptions of the teaching performance in implementing the quality assurance standards of the ELPP. They were graded on Likert's three-point scale, ranging from (high - moderate - low) and chosen for its simplicity and ease of understanding for respondents, particularly in contexts where more complex scales might lead to confusion. Additionally, this scale aligns with previous ELPP studies, supporting consistency in comparative analysis. One hundred sixty-two copies of the questionnaire were distributed with Arabic translation to ensure that the participants understand the items and respond honestly to them (Dörnyei, 2003). All copies of the questionnaire were returned, resulting in a 100% return rate. LikertLikert

The construction of the questionnaire underwent a structured process to verify its validity and reliability. Initially, it was sent to a panel of nine ELT specialists, professors, and quality assurance experts to assess its content validity, appropriateness, and clarity of each statement. Based on their feedback, the questionnaire was revised to better meet these standards. Subsequently, a pilot test was conducted to ensure the questionnaire practicality and effectiveness in capturing the necessary data. The reliability of the finalized version of the questionnaire was verified by calculating the Cronbach's Alpha Coefficient.

Table 2. Cronbach's Alpha Coefficient of the Questionnaire

Components	No. of items	Cronbach's Alpha Value
Teaching performance standards	22	0.771

The obtained Cronbach's alpha coefficient (0.771) indicated that the close-ended questionnaire exhibited good reliability. This level of reliability showed an acceptable internal consistency among the questionnaire items, indicating that the responses were consistently aligned with one another. In practical terms, a Cronbach's Alpha greater than 0.70 is generally deemed good, which reinforces the robustness of the data collected with this instrument (Dörnyei, Z. 2003; Dörnyei & Taguchi, 2010; George & Mallery, 2029).

3.3.Data Analysis

The first research question examined the students' perceptions of the implementation of indicators for teaching performance quality assurance standards. The participants' responses were quantitatively analyzed using means and standard deviations. Their responses were coded into Likert's three-point scale that corresponded to the following mean ranges: 1.00-1.66 (low), 1.67-2.33 (moderate), and 2.34-3.00 (high).

To answer the second research question, an independent sample T-test was computed to measure the differences between each pair of samples and identify discrepancies between the two variables of gender and students' level of study using SPSS, V. 22.

4. RESULTS OF THE STUDY

4.1. Results Related to the First Question of the Study

The first question of the study focused on students' perceptions of implementing indicators of teaching performance quality assurance standards in the ELPP.

Table 3. The Mean and Standard Deviation of the Students' Perceptions for Implementing Teaching Performance Standards

	Teachers' Performance	M	SD	Rank
1	The teacher specifies the course requirements to students: exams, readings, and worksheets at the beginning of the semester.	2.53	0.63	9
2	The teacher connects lecture topics to students' personal experiences.	2.38	0.70	15
3	The teacher concludes the lecture with a summary of the main points.	2.08	0.80	19
4	The teacher presents the topic in a logical and sequential manner.	2.60	0.56	4
5	The teacher demonstrates enthusiasm and energy during the lecture.	2.48	0.72	11
6	The teacher follows the curriculum in the sequence outlined in the course plan.	2.69	0.59	2
7	The teacher guides students on how to obtain supplementary educational resources (dictionaries, educational applications).	1.93	0.86	21
8	The teacher uses simple language and terminology.	2.55	0.61	8
9	The teacher assigns students educational activities (individual and group).	2.51	0.70	10
10	The teacher uses appropriate educational references (books).	2.17	0.72	18
11	The teacher employs educational technologies.	2.27	0.70	17
12	The teacher allows students to choose from assignments that will foster their creativity.	2.04	0.83	20
13	The teacher allocates part of the students' grades for class activities and participation.	2.55	0.68	7
14	The teacher employs various discussion methods, including class discussions and group discussions.	2.56	0.69	6
15	The teacher encourages students to think and be creative.	2.38	0.69	16
16	The teacher answers questions posed by students.	2.67	0.56	3
17	The teacher adheres to the scheduled start and end times for the lecture.	2.57	0.67	5
18	The teacher balances practical and theoretical aspects in the lecture.	2.41	0.68	14
19	The teacher grades students fairly according to their performance on the test.	2.76	0.50	1

20	The teacher assigns written homework related to the course content.	2.41	0.68	13
21	The teacher uses gestures and facial expressions during teaching.	2.45	0.69	12
22	The teacher allocates office hours for student consultations.	1.73	0.79	22
Total		2.40	0.32	

Table 3 presents the students' range of perceptions of teachers' implementation of the quality assurance standards for teaching performance indicators in the ELPP. The total mean score of all items was 2.40, and the standard deviation was 0.32. As Table 3 shows, the number of indicators (1,2,4,5,6,8,9, 13,14,15,16, 17,18,19,20,21) showed high means. Indicators number (3,7,10, 11, 12, 22) showed moderate mean perceptions. The results of the study showed that the students rated the indicators of teaching performance quality assurance standards as either highly or moderately implemented in the ELPP.

4.2.Results Related to the Second Question of the Study

The second question of the study investigated whether there is a statically significant difference attributed to the variables of the students' gender and their level of study in their mean responses regarding their perceptions of the implementation of quality assurance standards in teaching performance in the ELPP.

Table 4. T-Test Results for the Differences in the Students' Perceptions Based on the Variables of Gender and Level of Study.

Variable	M	SD	T- Value	df	P- Value
Male	2.34	0.31	3.84	160	0.000
Female	2.56	0.30			
First Level	2.42	0.33	1.08	160	0.28
Second Level	2.37	0.31			

Table 4 presents the independent sample T-test results of the students' perceptions of the implementation of the teaching performance indicators standards. As Table 4 indicates, the p-value reveals a statistically significant difference attributed to the gender variable at the significance level ($\alpha \leq 0.05$) in the indicators of teaching performance, as perceived by students in relation to the quality assurance standards. The difference favored the female students, with a mean of 2.56. However, the p-value shows that there was no statistically significant difference attributed to the students' level of study variable at the significance level of ($\alpha \leq 0.05$) in the indicators of teaching performance in light of the quality assurance standards from the students' perspectives.

5. DISCUSSION OF THE RESULTS

The findings of the first study question demonstrated that the indicators of teaching performance were highly to moderately implemented in the ELPP, based on the students' perceptions, as shown in Table 3. The students' positive perceptions of implementing the teaching performance indicators could be attributed to the fact that the teaching staff in the ELPP are highly qualified, holding MA and BA in English. Moreover, the ELPP administration requests that students assess the performance of the teaching staff annually, and they have never raised any complaints. The highly rated indicators of the teaching performance could also be

interpreted as fruitful results for the regularly conducted workshops for the teaching staff on various teaching issues.

The findings in the present study align with, and sometimes contrast with, previous research on evaluating teaching performance in educational institutions worldwide. Based on the findings in Table 4, the highly ranked indicators in this study are, the teachers "*grade students fairly*" (M=2.76, Rank 1), "*follow the curriculum sequence*" (M=2.69, Rank 2), and "*answer students' questions*" (M=2.67, Rank 3). These results align with Al-Mekhlafi (2020) and Marsh (2007) findings, which also identify instructional clarity and fair assessment as essential indicators of students' performance in evaluations. In line with the present results, Alsirr (2004) found that instructional planning scored highly (82%), reinforcing the importance of structured teaching. Indicators related to engagement strategies were highly rated, including teachers' "*use various discussion methods*" (M=2.56, Rank 6), "*encourage creativity*" (M=2.38, Rank 16), and "*connect topics to personal experiences*" (M=2.38, Rank 15). These results are congruent with Al-Rsa'i et al. (2024) findings, who found that trained teachers and female instructors excelled in interactive teaching. However, these results are inconsistent with Johari (2021) findings, which reported an average level of student engagement.

The findings regarding assessment indicators yield mixed results. Although "*fair grading*" was rated the highest (M=2.76) along with "*allocating grades for participation*" (M=2.55, Rank 7), "*offering creative assignment choices*" was rated relatively moderate (M=2.04, Rank 20). These findings mirror Canaan (2005) mixed results, where faculty met basic assessment standards but lacked innovative assessment strategies. Faraj and Al-Anqari (2020) also reported moderate adherence to quality standards in assessment, suggesting a need for more formative and student-centered evaluation methods.

Moreover, moderate teaching performance indicators are observed in "*concluding lectures with summaries*" (M = 2.08, Rank 19). This finding suggests a gap in reinforcing key concepts, which contrasts with Darling-Hammond (2017) emphasis on lesson consolidation as a best practice. Furthermore, the results indicate moderate teaching performance in teachers' "*employ educational technologies*" (M=2.27, Rank 17) and "*use appropriate educational references*" (M=2.17, Rank 18). These findings align with Zniati (2005) findings, as cited in Abu-Al-Sha'r and Al-Harashseh (2013), where faculty struggled with modern instructional tools despite institutional efforts toward accreditation.

Indicators related to guidance are rated moderately as the lowest in the present study, "*Allocating office hours for consultations*" (M=1.73, Rank 22). This result is concerning, as Abdulsalam and Al-Qataani (2024) found that student interaction skills were among the strongest aspects of faculty performance in their study. This finding suggests an institutional barrier to teacher accessibility, which may hinder student support. Teachers' inaccessibility might be attributed to the hour-based contracts that most teachers in the ELPP do not have office hours, as observed in the present study. In addition, "*guiding students to supplementary resources*" was rated as moderate (M = 1.93, Rank 21). This finding contradicts Darling-Hammond (2017) advocacy for resource accessibility as a critical teaching competency. This gap suggests a need for professional development in resource utilization. Teachers have to identify and locate relevant educational resources (e.g., textbooks, digital tools, online articles,

real-world examples, supplementary materials). They have to instruct students on how to utilize these resources for their learning, research, and skill development.

Table 4 shows that there is a statistically significant difference at the 0.05 Alpha level in the students' perceptions of implementing the teaching performance indicators attributed to the gender variable. In other words, female students rated the implementation of the teaching performance indicators higher than male students did. This finding might be interpreted as female students seem to be more satisfied with their female teachers' performance. They may exhibit more positive evaluation behaviors due to higher kindness (Costa et al., 2001) and social desirability bias, where women may avoid overly critical ratings to maintain harmony (Paulhus, 1991). Additionally, female students tend to have higher class attendance rates, which may lead to more positive perceptions of teaching effectiveness. Marsh et al. (2008) noted that female students often report a stronger relationship with teachers, particularly female instructors, which is also the case with the teachers of the girls in the ELPP. When most teachers in the study are women, female students might rate them more favorably due to shared gender identity (Dee, 2007). This finding is consistent with the results of Al-Rsa'i et al. (2024) and Abdulsalam and Al-Qataani (2024), who found that female students rated teachers more favorably than male students. Male students' lower rating of the implementation of the teaching performance indicators might be interpreted as they missed classes more frequently due to personal/family obligations, which could provide lower evaluations due to relatively reduced engagement (Rumberger, 2011).

The current study finds no significant differences between first and second-year students' perceptions. The absence of statistically significant differences between first- and second-year students indicates that both cohorts perceive the quality of teaching performance in a relatively similar manner. This uniformity may reflect the program's consistent instructional delivery and alignment with quality assurance standards across different academic grade levels. It also implies that students' level of study does not substantially influence their evaluation criteria or learning expectations within the ELPP.

6. CONCLUSION AND IMPLICATIONS

To enhance teaching effectiveness in the ELPP, comprehensive professional development is essential. This development should be categorized based on resource demand and potential impact: high, medium, and low. High-resource-demand interventions includes revising the hour-based contract policy of the ELPP teachers. Al-Ahgaff University may need to consider offering yearly-based contracts or tenured positions, that require teachers to allocate office hours for student consultations similarly to the tenured professors at Al-Ahgaff University. Furthermore, medium-resource-demand interventions might include comprehensive digital pedagogy training, which emphasizes the effective integration of technology for interactive lessons, collaborative activities, and differentiated instruction—necessitating sustained access to relevant software and continuous technical support. In addition, medium-resource-demand interventions might involve workshops on developing more creative assessment approaches and advanced lesson techniques, such as interactive summaries, concept mapping, and robust Q&A sessions, to ensure knowledge consolidation. Low-resource demand interventions could focus on enhancing resource literacy, enabling instructors to strategically identify, evaluate,

and integrate diverse educational resources, including online dictionaries, language learning applications, and authentic English materials. The Ministry of Higher Education may benefit from the findings of this study in establishing teaching performance indicators for quality assurance standards and accreditation, as well as in evaluating English language proficiency programs in public Yemeni universities. The Ministry needs to conduct further symposia and invite pedagogists to discuss effective teaching and performance indicators in English language proficiency programs. Adhering to quality standards in higher education is vital for enhancing institutional effectiveness, academic credibility, and stakeholder trust. Quality assurance standards ensures systematic improvement in teaching, curriculum, and student services while promoting transparency, global alignment, and a culture of accountability and continuous development (Campbell & Rozsnyai, 2002; ENQA, 2015; Harvey & Green, 1993; Vlasceanu et al., 2005).

6.1.Limitations and Recommendations for Future Studies

The findings of the present study should be interpreted in light of its primary goal. The results were based on students' perceptions, which might be biased, to evaluate the implementation of teaching performance indicators. Further studies may incorporate other methods to evaluate the teaching performance, such as classroom observation. Additionally, the number of girls in the ELPP should be considered when interpreting the study findings. Their number was proportionally smaller compared to the boys. Accordingly, their smaller representation in the present study might have impacted the results, but it was an inevitable case.

REFERENCES

- Abdulsalam, G. M., & Al-Qataani, D. W. (2024). Evaluating the effectiveness of faculty members' performance in light of quality standards from the perspective of Arts College students at Ajdabiya University. *Journal of Educational Sciences*, <https://educational-sciences-journal.lacts.org.ly/esj/index.php/esj/article/view/449>
- Abu-Al-Sha'r, A. M., & Al-Harabsheh, M. A. (2013). Criteria of total quality management of faculty teaching skills: Perceptions of university students. *International Research in Education*, 1(1), 101–115. <https://doi.org/10.5296/ire.v1i1.3685>
- Aimah, S., & Purwanto, B. (2019). Evaluating teachers' performance: A need for effective teaching. *CELT: A Journal of Culture, English Language Teaching & Literature*, 19(1), 158–170. <https://doi.org/10.24167/celt.v19i1.1369>
- Airasian, P. W. (2000). *Assessment in the classroom: A concise approach* (2nd ed.). Boston: McGraw-Hill.
- Alaidaros, H., Kherd, A., & Al-Aidroos, H. A. (2021). Students' perceptions of online learning during Covid-19 pandemic at Al-Ahgaff University, Yemen: A Survey. In *the 2021 International Conference of Modern Trends in Information and Communication Technology Industry (MTICTI)* (pp. 1-5). IEEE. Retrieved from https://www.researchgate.net/publication/357592415_Students%27_Perception_of_Online_Learning_during_Covid-19_Pandemic_at_Al-Ahgaff_University_Yemen_A_Survey

- Alfalagg, A. (2018). *Impact of teacher-student writing conferences on cohesion in EFL students' writing: An experimental study*. [Unpublished doctoral dissertation, The English and Foreign Languages University]. Hyderabad, India.
- Alfalagg, A., & Ba-Udhan, H. (2023). Teachers' perceptions on using Arabic L1 to teach English in higher education in Yemen. In E. Meletiadou (Ed.), *Handbook of research on implications of sustainable development in higher education* (pp. 251–271). IGI Global. <https://doi.org/10.4018/978-1-6684-6172-3.ch012>
- Al-Gahali, M. A., Al-Jarban, N. S., & Al-Banna, K. D. (2022). Availability of programmatic academic accreditation standards in academic programs at Amran University. *Amran University Journal*, 20, 217–46. <https://doi.org/10.59145/jaust.v2i3.33>
- Al-Hakmi, I. H. (2004). Professional competencies required for university professors from their students' perspectives and their relationship with some variables. *Journal of the Message of the Arab Gulf*, 24(90), 13–56. <https://doi.org/10.35270/0011-024-090-001>
- Ali, M., & Shastri, R. K. (2010). Implementation of total quality management in higher education. *Asian Journal of Business Management*, 2(1), 9–16. <https://maxwellsci.com/print/ajbm/v2-9-16.pdf>
- Alkyumi, A. (2002). *Estimating the degree of applicability of certain concepts of total quality management from the administrators and teachers' perspectives in the Faculty of Education in the Sultanate of Oman*. [Unpublished MA Thesis], Yarmouk University, Irbid - Jordan.
- Al-Maraziq, I. A., & Bani Ahmed, O. M. (2022). Assessing the quality of academic programs in the College of Educational Sciences at Jerash University from the students' point of view. *Humanitarian & Natural Sciences Journal*, 3(1), 662–678. <https://doi.org/10.53796/hnsj3140>
- Al-Mekhlafi, M. S. (2020). Quality evaluation of postgraduate programs from the perspective of students (Faculty of Education, Imam Abdulrahman Bin Faisal University (IAU), KSA). *Journal of Education in Black Sea Region*, 5(2), 76–95. <https://doi.org/10.31578/jeb.v5i2.201>
- Almuslimi, F. Kh. (2011). *Evaluating the English language teacher professional preparation program in the Faculty of Education, Sana'a University, in the light of quality assurance standards* [Unpublished doctoral dissertation]. Sana'a University. <https://doi.org/10.26822/iejee.2024.332>
- Alotumi, M. (2015). *The use of CALL for EFL instruction in ELTE programmes in Yemeni universities: An impact study*. [Unpublished doctoral dissertation, The English and Foreign Languages University]. Hyderabad, India.
- Al-Rsa'i, M. S., Waswas, D., Altawarah, A., & Al-Rowad, F. (2024). Evaluating the Quality of Teaching Performance among Jordanian Teachers in Light of Certain Demographic Variables. *International Electronic Journal of Elementary Education*, 16(3), 295–309. <https://doi.org/10.26822/iejee.2024.332>
- Al-Saba, S. S., Ghalib, A. H., & Abdo, S. A. (2010). Evaluating the Arabic language teacher preparation program in the College of Education at Sana'a University in the light of

- comprehensive quality standards. *Arab Journal of Quality Assurance of University Education*, 5, 96-130. Retrieved from <https://search.shamaa.org/fullrecord?ID=37310>
- Al-Sirr, K. K. (2004). Evaluation of the quality of university teaching skills among faculty members at Al-Aqsa University in Gaza: From their perspective and its relationship with some variables. *Journal of Al-Aqsa University – Humanities Series*, 8(2), 274–315. <https://search.emarefa.net/detail/BIM-245798>
- Al-Subaie, M. H. (2010). The reality of female faculty members' teaching skills in applied sciences in the light of the comprehensive quality standards from the students' perspectives at Umm Al-Qura University. Seminar of higher education for girls: Dimensions and outlooks. Tiba University in Saudi Arabia. Retrieved from <https://search.mandumah.com/Record/100474>
- Ameyaw, D. A., & Khumalo, S. S. (2024). The involvement, engagement, and satisfaction of higher education students: a literature review. *International Journal of Religion*, 5(9), 638–647. <https://doi.org/10.61707/fr1e9j19>
- Arnon, S., & Reichel, N. (2007). Who is the ideal teacher? Am I? Similarity and differences in perception of students of education regarding ideal and actual self. *Teachers and Teaching: Theory and Practice*, 13(5), 441–464. <https://doi.org/10.1080/13540600701561653>
- Ateeq, E. (2023). *Evaluating the academic program of the English Language Department at College of Women, Seiyun University in the light of quality standards and national academic accreditation* [Unpublished master's thesis]. Seiyun University.
- Ba-udhan, H. (2014). *Investigating the problems and strategies of listening comprehension of EFL student-teachers: A study in the Yemeni context*. [Unpublished doctoral dissertation, The English and Foreign Languages University]. Hyderabad, India.
- Berk, R. A. (2018). *Top 10 strategies for leveraging student feedback to improve teaching*. Baltimore: Johns Hopkins University Press.
- Bi, S.-Y. (2012). Evaluation of teaching effectiveness in Taiwan's EFL context. In *Selected papers from the twenty-first international symposium and book fair on English teaching* (pp. 162–170). Retrieved from <https://www.ltc.ntu.edu.tw/files/20221028160822203.pdf>
- British Council. (2018). *Quality standards in English language teaching*. Retrieved from <https://www.britishcouncil.org/>
- Brown, J. D. (1989). Language program evaluation: A synthesis of existing possibilities. In R. K. Johnson (Ed.), *The second language curriculum* (pp. 222–241). Cambridge: Cambridge University Press.
- Brown, J. D. (1995). *The elements of language curriculum: A systematic approach to program development*. Boston: Heinle & Heinle Publishers.
- Campbell, C., & Rozsnyai, C. (2002). *Quality assurance and the development of course programmes*. Bucharest: UNESCO-CEPES.
- Canaan, A. (2005). *Improving faculty members' performance according to total quality criteria, the twelfth annual National Conference (Al-Arabi IV)*. Improving the

- performance of Arab Universities in the light of the standards of quality management and accreditation systems. 18-19 December. Retrieved from <https://abhathna.com/files/maqa/1181.pdf>
- Carless, D. (2019). Feedback loops and the longer-term: Towards feedback spirals. *Assessment & Evaluation in Higher Education*, 44(5), 705–714. <https://doi.org/10.1080/02602938.2018.1531108>
- Costa, P. T., Jr., Terracciano, A., & McCrae, R. R. (2001). Gender differences in personality traits across cultures: Robust and surprising findings. *Journal of Personality and Social Psychology*, 81(2), 322–331. <https://doi.org/10.1037/0022-3514.81.2.322>
- Council of Europe (2020), Common European framework of reference for languages: learning, teaching, assessment – Companion volume, Strasbourg: Council of Europe Publishing, Retrieved from: www.coe.int/lang-cefr.
- Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). Thousand Oaks: SAGE Publications.
- Crosthwaite, P. (2016). A longitudinal multidimensional analysis of EAP writing: Determining EAP course effectiveness. *Journal of English for Academic Purposes*, 22, 166–178. <https://doi.org/10.1016/j.jeap.2016.04.005>
- Crystal, D. (2012). *English as a global language* (2nd ed.). NY: Cambridge University Press.
- Darling-Hammond, L. (2017). Teacher education around the world: What can we learn from international practice? *European Journal of Teacher Education*, 40(3), 291–309. <https://doi.org/10.1080/02619768.2017.1315399>
- Darling-Hammond, L., Newton, S. P., & Wei, R. C. (2013). Developing and assessing beginning teacher effectiveness: The potential of performance assessments. *Educational Assessment, Evaluation and Accountability*, 25(3), 179–204. <https://doi.org/10.1007/s11092-013-9163-0>
- Dee, T. S. (2007). Teachers and the gender gaps in student achievement. *Journal of Human Resources*, 42(3), 528–554. <https://doi.org/10.3368/jhr.XLII.3.528>
- Dörnyei, Z. (2003). *Questionnaires in second language research: Construction, administration, and processing*. Mahwah: Lawrence Erlbaum
- Dörnyei, Z., & Taguchi, T. (2010). *Questionnaires in Second language research: Construction, administration and processing*, (2nd ed.). NY: Routledge.
- English Language Unit. (2016). *ELPP outline* [Unpublished internal document]. Al-Ahgaff University.
- European Association for quality assurance in higher education area (ESG). (2015). Brussels, Belgium. Retrieved from https://enqa.eu/wp-content/uploads/2015/11/ESG_2015.pdf
- Faraj, K. M., & Al-Anqari, M. A. (2020). The availability of quality standards in teaching performance of faculty members at Gulf of Sidra University from students' perspective, *Journal of Shart University for Humanities*, 50-76. <https://doi.org/10.37375/sujh.v10i2>
- Garrett, R., & Steinberg, M. P. (2015). Examining teacher effectiveness using classroom observation scores: Evidence from the randomization of teachers to students. *Educational Evaluation and Policy Analysis*, 37(2), 224–242. <https://doi.org/10.3102/0162373714537551>

- George, D., & Mallery, P. (2019). *IBM SPSS statistics 26 step by step: A simple guide and reference (16th ed.)*. NY: Routledge.
- Harris, J. (2009). Late-stage refocusing of Irish-language programme evaluation: Maximizing the potential for productive debate and remediation. *Language Teaching Research*, 13(1), 55-76. <https://doi.org/10.1177/1362168808095523>
- Harvey, L., & Green, D. (1993). Defining quality. *Assessment & Evaluation in Higher Education*, 18(1), 9–34. <https://doi.org/10.1080/0260293930180102>
- Hattie, J. A. C. (2012). *Visible learning for teachers: Maximizing Impact on Learning* (1st ed.). London: Routledge. <https://doi.org/10.4324/9780203181522>
- Jaber, S. S. (2021). *Evaluating the academic program of the English language department at Women's College, Hadhramout University*. [Doctoral dissertation]. Hadhramout University.
- Johari, S. (2021). Evaluating university teaching performance in light of education quality assurance standards from students' perspectives. *Revue des Sciences Humaines et Sociales*, 7(2), 361-391. <https://asjp.cerist.dz/en/article/155709>
- Keefe, K., & Shi, L. (2017). *An EAP program and students' success at a Canadian university*. *TESL Canada Journal*, 34(2), 1–24. <https://doi.org/10.18806/tesl.v34i2.1264>
- Koslowski, F. A. (2006). Quality and assessment in context: A brief review. *Quality Assurance in Education*, 14(3), 277–288. <https://doi.org/10.1108/09684880610678586>
- Lynch, B. K. (1996). *Language program evaluation: Theory and practice*. Cambridge: Cambridge University Press.
- Marsh, H. W. (1987). Students' evaluations of university teaching: Research findings, methodological issues, and directions for future research. *International Journal of Educational Research*, 11(3), 253–388. [https://doi.org/10.1016/0883-0355\(87\)90001-2](https://doi.org/10.1016/0883-0355(87)90001-2)
- Marsh, H. W. (1987). Students' evaluations of university teaching: Research findings, methodological issues, and directions for future research. *International Journal of Educational Research*, 11(3), 253–388. [https://doi.org/10.1016/0883-0355\(87\)90001-2](https://doi.org/10.1016/0883-0355(87)90001-2)
- Marsh, H. W. (2007). Students' evaluations of university teaching: A multidimensional perspective. In R. P. Perry & J. C. Smart (Eds.), *The scholarship of teaching and learning in higher education: An evidence-based perspective* (pp. 319–384). Netherlands: Springer
- Marsh, H. W., Martin, A. J., & Cheng, J. H. S. (2008). A multilevel perspective on gender in classroom motivation and climate: Potential benefits of male teachers for boys? *Journal of Educational Psychology*, 100(1), 78–95. <https://doi.org/10.1037/0022-0663.100.1.78>
- Norris, J. M. (2009). Understanding and improving language education through program evaluation: Introduction to the special issue. *Language Teaching Research*, 13(1), 7–13. <https://doi.org/10.1177/1362168808095520>
- Owen, J. M. (2007). *Program evaluation: Forms and approaches*. New York: Guilford Press.
- Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of personality and social psychological*

- attitudes (pp. 17–59). San Diego: Academic Press. <https://doi.org/10.1016/B978-0-12-590241-0.50006-X>
- Peacock, M. (2009). The evaluation of foreign-language-teacher education programmes. *Language Teaching Research*, 13(3), 259–278. <https://doi.org/10.1177/1362168809104698>
- Penny, A. R., & Coe, R. (2004). Effectiveness of consultation on student ratings feedback: A meta-analysis. *Review of Educational Research*, 74(2), 215–253. <https://doi.org/10.3102/00346543074002215>
- Petersen, J. C. (1999). *Internationalizing quality assurance in higher education*. Washington: Council for Higher Education Accreditation.
- Ramsden, P. (1991). A performance indicator of teaching quality in higher education: The course experience questionnaire. *Studies in Higher Education*, 16(2), pp. 129–150. doi:<http://dx.doi.org/10.1080/03075079112331382944>
- Rasool, S. (2010). *A comparative study of the quality assurance practices in public and private universities* [Unpublished doctoral dissertation]. University of the Punjab, Pakistan.
- Redder, L.A. (2010). Quality assurance in Redder, L. A. (2010). *Quality assurance in professional higher education institutions: The structure of internal quality assurance systems and the experienced costs* (Master's thesis, University of Twente).
- Richards, J. C. (2017). *Curriculum development in language teaching* (2nd ed.). Cambridge: Cambridge University Press. <https://doi.org/10.1017/9781009024556>
- Richards, J. C. (2017). Teaching English through English: Proficiency, pedagogy, and performance. *RELC Journal*, 48(1), 7–30. <https://doi.org/10.1177/0033688217690059>
- Richards, J. C., & Farrell, T. S. C. (2005). *Professional development for language teachers: Strategies for teacher learning*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511667237>
- Richards, J. C., & Farrell, T. S. C. (2005). *Professional development for language teachers*. Cambridge: Cambridge University Press.
- Santiago, P., Roseveare, D., van Amelsvoort, G., Manzi, J., & Matthews, P. (2009). *Teacher evaluation in Portugal*: Paris: Organization for Economic Co-operation and Development. Retrieved from <https://gpseducation.oecd.org/Content/ProjectsMaterial/EvaluationAndAssessment/TeacherEvaluationInPortugal.pdf>
- Seidel, T., & Shavelson, R. J. (2007). Teaching effectiveness research in the past decade: The role of theory and research design in disentangling meta-analysis results. *Review of Educational Research*, 77(4), 454–499. <https://doi.org/10.3102/0034654307310317>
- Spooren, P., Brockx, B., & Mortelmans, D. (2013). On the validity of student evaluation of teaching: The state of the art. *Review of Educational Research*, 83(4), 598–642. <https://doi.org/10.3102/0034654313496870>
- Spooren, P., Brockx, B., & Mortelmans, D. (2017). On the validity of student evaluation of teaching: The state of the art. *Review of Educational Research*, 87(1), 143–186. <https://doi.org/10.3102/0034654316663798>
- Storch, N., & Tapper, J. (2009). The impact of an EAP course on postgraduate writing. *Journal of English for Academic Purposes*, 8(3), 207–223. <https://doi.org/10.1016/j.jeap.2009.03.001>

- Surujlal, J. (2014). Improving service quality in our classrooms: Student feedback as a quality mechanism for teaching. *Mediterranean Journal of Social Sciences*, 5(13), 191–197. <https://doi.org/10.5901/mjss.2014.v5n13p191>
- Syahrudin, B., Syamsudduha, & Haryanto. (2013). Teachers' pedagogical competence in school-based management. *Journal of Education and Learning*, 7(4), 213–218. <https://doi.org/10.11591/edulearn.v7i4.18491>
- TESOL International Association. (2019). Standards for ESL/EFL teachers of adults. Retrieved from <https://www.tesol.org>
- UNICEF. (2023). *Yemen education in emergencies report*. Retrieved from <https://open.unicef.org/sites/transparency/files/2024-05/Yemen%20CER%202023.pdf>
- Usmani, M. A. W., Khatoon, S., Shammot, M. M., & Zamil, A. M. (2012). Towards a network of quality assurance in higher education: A Pakistani model. *Archives des Sciences*, 65(7), 224–229. <https://www1.duhs.edu.pk/app-assets/filemanager/files/Towards%20a%20Network%20of%20Quality%20Assurance%20in%20Higher%20Education%20A%20Pakistani%20Model.pdf>
- Uzun, L. (2016). Evaluation of the latest English language teacher training program in Turkey: Teacher trainees' perspective. *Cogent Education*, 3(1), 1147115. <https://doi.org/10.1080/2331186X.2016.1147115>
- Vlăsceanu, L., Grünberg, L., & Pârlea, D. (2005). *Quality assurance and accreditation: A glossary of basic terms and definitions*. Bucharest: Unesco-Cepes.
- Vlăsceanu, L., Grünberg, L. and Pârlea, D. (2007). *Quality assurance and accreditation: A glossary of basic terms and definitions*. Bucharest: UNESCO-CEPES. https://www.observatoriorh.org/sites/default/files/webfiles/fulltext/unesco_cepes_qual_assur_acred.pdf
- Vroeijenstijn, A., L., (1995). *Improvement and accountability: Navigating between Scylla and Charybdis*. London, Bristol & Pennsylvania: Jessica Kinsley
- Watty, K. (2006). Addressing the basics: Academics' view of the purpose of higher education. *The Australian Educational Researcher*, 33(1), 23–39. <https://doi.org/10.1007/BF03246279>