International Journal of Evaluation and Research in Education (IJERE)
Vol. 15, No. 3, June 2026, pp. 1971~1985
ISSN: 2252-8822, DOI: 10.11591/ijere.v15i3.37326

Validating the factor structure of primary school teaching quality using the PDCA cycle

Nhat Thong Du[1], Van Dat Tran[2]
[1]Faculty of Educational Sciences, Ho Chi Minh City University of Education, Ho Chi Minh City, Vietnam
[2]Faculty of Education, An Giang University, Vietnam National University Ho Chi Minh City, Long Xuyen, Vietnam

Article Info

Article history:

Received Oct 16, 2025
Revised Mar 3, 2026
Accepted Mar 28, 2026

Keywords:

PDCA cycle
Primary education
Scale validation
Teaching quality
Vietnam

ABSTRACT

Improving teaching quality is a critical goal of educational reform, particularly in the context of Vietnam's 2018 General Education Program (GEP). However, there is still a lack of valid instruments for assessing instructional practices based on continuous improvement models. This research fills this void by creating and validating the primary school teaching quality scale (PSTQS), based on the Plan-Do-Check-Act (PDCA) cycle, a well-known quality management framework. A cross-sectional survey design was used to collect data from 528 primary school teachers in Ho Chi Minh City. The 20-item PSTQS was constructed to align with the PDCA model and the competencies outlined in the 2018 GEP. The sample was randomly split into two groups for a two-phase validation: exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). EFA found a four-factor structure that matched the PDCA cycle and accounted for 45.3% of the variance. CFA confirmed the model's fit (comparative fit index (CFI)=0.95, Tucker-Lewis index (TLI)=0.94, root mean square error of approximation (RMSEA)=0.04, standardized root mean square residual (SRMR)=0.05). The scale demonstrated strong internal consistency (composite reliability (CR)>0.79), with robust convergent and discriminant validity. The PSTQS is a reliable, valid tool for evaluating teaching practices through a quality improvement lens. While results are specific to Ho Chi Minh City, the scale offers a foundation for broader application and supports continuous professional development and policy implementation.

This is an open access article under the CC BY-SA license.

Corresponding Author:

Van Dat Tran
Faculty of Education, An Giang University, Vietnam National University Ho Chi Minh City
Long Xuyen, An Giang, Vietnam
Email: tvdat@agu.edu.vn

1. INTRODUCTION

In the context of global educational reform, Vietnam initiated a major transformation in 2018 with the introduction of its new General Education Program (GEP), promulgated by the Ministry of Education and Training (MOET) [1]. At the primary level, the GEP aims to promote holistic development by fostering harmonious physical, intellectual, and moral growth. It emphasizes five core qualities—patriotism, kindness, diligence, honesty, and responsibility—alongside ten essential competencies spanning general and subject-specific domains. These goals have been operationalized into measurable learning outcomes across grade levels and subjects, requiring primary teachers not only to deliver content effectively but also to systematically plan, assess, and refine instruction in alignment with national standards. Amid rising global demands for accountability and continuous improvement, many education systems have adopted quality management approaches. The total quality management (TQM), with its emphasis on learner-centered

processes, feedback mechanisms, and stakeholder engagement, has gained recognition as a framework for strengthening school effectiveness and learning outcomes [2]–[6]. Research identifies critical success factors (CSFs)—including leadership commitment, participatory culture, and data-informed decision-making—as essential for successful TQM implementation across diverse contexts [7]. In Vietnam, initial efforts to apply TQM principles suggest potential benefits for instructional quality; however, implementation remains fragmented, particularly in primary education, where systemic integration has yet to be fully realized.

Within the TQM framework, the Plan-Do-Check-Act (PDCA) cycle serves as a central mechanism for continuous improvement. Widely embedded in the internal quality assurance systems (IQAS), the PDCA has demonstrated effectiveness in enhancing teacher competencies, curriculum alignment, and feedback systems across international settings [8], [9]. Empirical studies report its contributions to improving digital literacy among primary teachers [10], aligning pedagogy with accreditation standards [11], and supporting instructional reform in both Southeast Asian and Western contexts. In Vietnam, researchers have suggested the PDCA as a viable instrument for tackling instructional difficulties within the new GEP [12], [13]. Nevertheless, existing work has largely remained conceptual or focused on higher education, leaving the systematic application of the PDCA to primary classroom practice underdeveloped. Despite growing recognition of the PDCA's value, a critical gap persists: the absence of a standardized and psychometrically validated instrument for measuring PDCA implementation in Vietnamese primary school teaching.

Although conceptual models and localized pilots have been suggested, no empirically tested scale currently captures how teachers design, implement, evaluate, and improve instruction within a PDCA framework. International scholarship has increasingly developed and validated the TQM-based educational quality instruments in countries such as the United Kingdom, Japan, Malaysia, and Spain [14]–[16]. However, these tools reflect contextual features—such as institutional autonomy, governance structures, and technological infrastructure—that differ substantially from Vietnam's centralized education system. Consequently, direct adoption is inappropriate, indicating the importance of a culturally grounded and psychometrically robust instrument.

To fill this gap, the current study aims: i) to establish and test the four-factor PDCA structure of the primary school teaching quality scale (PSTQS) as distinct yet interconnected dimensions of teaching quality, and ii) to evaluate the construct validity and reliability of the PSTQS by conducting internal consistency, convergent validity, and discriminant validity tests. Based on the research objectives, the present study will tests two hypotheses:

- H1: the PSTQS comprises four distinct but interrelated factors corresponding to teaching design (Plan), teaching implementation (Do), teaching evaluation (Check), and teaching improvement (Act).
- H2: the PSTQS demonstrates strong psychometric properties, including high internal consistency and acceptable convergent and discriminant validity.

This research advances TQM theory by implementing the PDCA cycle within the classroom, converting it from an organizational quality framework into a pedagogically based model of instructional practice. It enhances continuous improvement theory by empirically substantiating teaching quality as a cyclical, multidimensional construct (PDCA). This study enhances educational quality assurance theory by integrating PDCA logic with competency-based reform under Vietnam's 2018 GEP, utilizing a context-sensitive, psychometrically validated measurement model. Ultimately, it fortifies the basis for sustainable quality improvement in Vietnamese primary education.

2. LITERATURE REVIEW

2.1. Total quality management in the educational context

In recent decades, education systems worldwide have transitioned to a more transparent, accountable, and perpetually improving approach. It is anticipated that schools at all levels will adhere to national benchmarks and international standards while also accommodating the changing expectations of stakeholders [17]. TQM has emerged as a comprehensive framework for institutional development within this reform landscape. Its core principles, such as systematic planning, stakeholder engagement, continuous improvement, and data-driven decision-making, closely align with contemporary educational priorities. Educational institutions have consistently emphasized the importance of leadership in the development of a quality culture, staff engagement, and strategic direction, adapting established quality models such as the European Foundation for Quality Management (EFQM) [18]. Research shows a positive correlation between the implementation of TQM and student achievement, teacher performance, and school effectiveness [19], [20]. In particular, the quality of the teacher remains a critical factor in the learning outcomes of students [4]. The quality assurance initiatives in general education in Vietnam reflect this global trend. Despite the potential of early attempts to incorporate TQM principles, their scope and sustainability are still restricted [7], [12]. The absence of standardized measurement instruments, resistance to change, and resource

constraints are common challenges [6]. These obstacles emphasize the necessity of identifying context-specific success factors and creating dependable tools to institutionalize quality improvement and support evidence-based reform.

2.2. PDCA cycle as a mechanism for continuous improvement in teaching

Within the TQM framework, the PDCA cycle serves as a fundamental mechanism for ongoing enhancement in education [21]–[23]. The cycle structures instructional enhancement into four iterative phases: i) establish objectives and formulate strategies that correspond with desired learning outcomes; ii) execute instructional plans within classroom practices; iii) assess: analyze educational outcomes in relation to established objectives; and iv) modify: adapt teaching methods based on reflective evaluation. This cyclical model fosters systematic reflection, instructional coherence, and adaptability to student requirements. The significance is particularly apparent in Vietnam's 2018 GEP, which prioritizes competency-based, student-centered learning. The GEP necessitates instruction that is purpose-driven, measurable, and adaptable—attributes inherently facilitated by the PDCA methodology. Research indicates that the PDCA is crucial for fortifying internal quality assurance [9], improving teacher competencies [10], and facilitating accreditation systems [11]. Scholars in Vietnam have proposed the adoption of the PDCA framework to improve subject-specific instruction, particularly in mathematics and science [12]. However, practical implementation is often inconsistent and frequently lacks structured diagnostic feedback mechanisms [13]. These challenges highlight the need for a reliable tool that can assess how well teachers are using PDCA principles in their everyday work, which would help improve their skills and ensure accountability in the GEP.

2.3. Importance of scale validation in educational research

In order for theoretical models like the PDCA to effectively inform practice, it is imperative to have psychometrically sound measurement instruments. A scale that is considered valid must exhibit construct clarity, reliability, and empirical robustness. Exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) are frequently implemented as complementary methodologies for verifying internal structure in educational research. In the initial phases of scale development, EFA is commonly utilized to uncover latent dimensions without enforcing a predefined structure [16], [23]. The proposed model's goodness-of-fit is subsequently assessed by CFA, which confirms construct validity through empirical verification [16], [24]. In the context of educational quality assurance and TQM, numerous studies have successfully utilized EFA and CFA to validate instruments that measure organizational performance and instructional quality [25], [26]. Researchers are increasingly implementing a two-sample cross-validation strategy, which involves conducting EFA and CFA on independent samples, to improve generalizability and mitigate sample bias [27]. In the context of Vietnam's educational reform, where the PDCA is becoming more deeply ingrained in instructional design, it is both methodologically necessary and practically urgent to create a reliable and context-sensitive scale. A scientific foundation would be established for teachers and policymakers to monitor and sustain instructional improvement through the use of such an instrument.

3. METHOD

3.1. Research framework

This study adopts a quantitative, cross-sectional survey design to develop and validate an instrument for assessing teaching quality in primary schools. The research is theoretically grounded in the PDCA cycle, a foundational model in continuous quality improvement. Within the educational context, the PDCA framework offers a structured, cyclical approach that supports teachers in systematically enhancing their instructional practices over time. Extant literature highlights the relevance of the PDCA model in guiding teaching-related processes. Specifically, it outlines a continuous loop in which educators: i) establish instructional objectives and design learning plans (Plan); ii) implement these strategies in the classroom (Do); iii) evaluate outcomes and assess student learning (Check); and iv) use findings to inform changes and enhance future instruction (Act) [8], [13].

The present study operationalizes the PDCA cycle into four sequential and interrelated dimensions that reflect the key stages of managing teaching quality in primary education: i) Plan—teaching design, ii) Do—teaching implementation, iii) Check—teaching evaluation, and iv) Act—teaching improvement, drawing on this framework. This conceptual framework posits that effective teaching is an iterative and evidence-based process, as illustrated in Figure 1, rather than a linear or one-time effort. The model illustrates the importance of continuous professional development, data-informed decision-making, and reflective practice by organizing the teaching process around the PDCA cycle. This theoretical foundation not only informs the development of the measurement instrument but also aligns with the objectives of competency-based instruction as delineated in Vietnam's 2018 GEP.

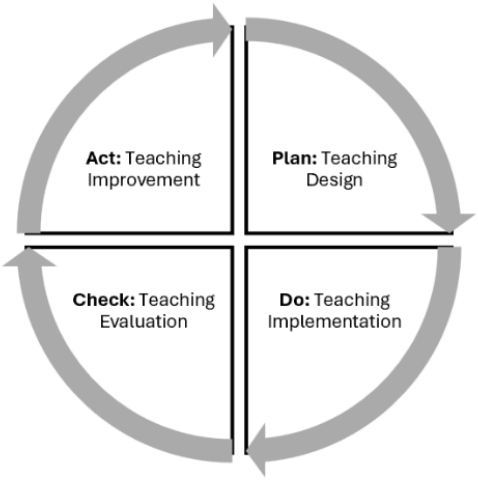

Figure 1. Conceptual framework of the PDCA cycle for teaching quality management

3.2. Participants

In-service teachers from public primary schools in Ho Chi Minh City, Vietnam, comprised the participants of this investigation. The study implemented a convenience sampling approach to efficiently recruit participants from a diverse array of institutions and geographic regions. An online survey was distributed through Google Forms to teachers, who were invited to participate. This method was chosen for its practicality, cost-effectiveness, and capacity to engage a substantial number of participants in a relatively short period of time, which is a critical factor in exploratory and confirmatory psychometric research. A total of 528 valid responses were obtained. Although convenience sampling is beneficial for the initial development and validation of scales, it does have limitations in terms of generalizability. In particular, the results may not accurately reflect the broader population of primary school teachers in Vietnam, particularly those in more remote or underserved regions, as participants self-selected into the study and were recruited from a single metropolitan area. Nevertheless, the sample size and the diversity of demographic characteristics serve to alleviate some of these apprehensions. In accordance with the two-sample cross-validation strategy proposed by MacCallum *et al.* [27], the total sample was randomly divided into two independent subgroups to guarantee methodological rigor during the scale validation process. By reducing sample-specific bias and testing the replicability of the factor structure, this method improves the robustness of the findings. The underlying factor structure was uncovered by employing the first subgroup (n=281) for EFA. The proposed model was validated using the second subgroup (n=247) for CFA. The final sample of 528 teachers satisfies the commonly accepted standards for factor analysis in terms of sample size adequacy. This sample is deemed statistically sufficient for both EFA and CFA in accordance with conventional guidelines (e.g., a minimum of 200 cases per analysis or 5–10 participants per item) [27]. Furthermore, power analyses indicate that sample sizes exceeding 200 yield reliable parameter estimates and exceptional model fit evaluations in structural equation modeling scenarios.

A breakdown of demographic characteristics for both subsamples is provided in Table 1. As shown, the random assignment yielded two demographically comparable groups across gender, education level, teaching experience, school type, and location, thereby supporting the validity of the cross–validation approach. The study strictly followed all ethical rules. All participants were informed about the research purpose, participated voluntarily, and provided informed consent. Responses were collected anonymously, and data were handled with confidentiality in compliance with ethical standards for educational research.

3.3. Research instrument

The PSTQS was created through a multi-phase, structured, theory-driven process that was designed to guarantee conceptual rigor and contextual alignment with Vietnam's 2018 GEP. The generation of items was explicitly based on two complementary theoretical foundations. The 2018 GEP served as the initial foundation, outlining the professional responsibilities, formative assessment principles, and competency-based learning outcomes that primary school teachers are expected to fulfill [1]. The second foundation was the PDCA cycle, a continuous improvement framework that is widely used in instructional reform and educational quality management [8], [13]. The instrument was developed as a context-sensitive measure that reflects the instructional practices necessary for the effective implementation of the GEP, rather than merely

as a general teaching evaluation tool, by incorporating these two perspectives. The development process concentrated on operationalizing PDCA stages as observable instructional behaviors, rather than directly replicating statements from policy documents or management theory.

The transition from theory to measurement involved three conceptual steps. Initially, the four PDCA dimensions were reformulated in educational terms: the plan was conceptualized as instructional preparation and goal alignment; pedagogical execution and classroom implementation; checking as monitoring and formative assessment; and acting as professional growth and reflective refinement. This reinterpretation guaranteed that the quality management framework could be effectively implemented in classroom-level practice. Secondly, the PDCA framework was used to extract and align the key behavioral expectations that are embedded in the GEP. The GEP draws attention to the alignment of lesson objectives with competency standards, the implementation of student-centered teaching methods, the ongoing use of formative assessment, and the sustained reflection and professional development of teachers [1]. In the context of Vietnamese primary education, these reform principles offered substantive guidance for defining the effective PDCA implementation. Third, concrete, behavior-based items were developed by translating theoretical principles. To prevent the use of abstract managerial language, each item was composed to describe a specific, teacher-controlled action that reflects a competency explicitly or implicitly required by the GEP and clearly corresponds to one PDCA stage.

Table 1. Demographic characteristics of the EFA and CFA samples

Characteristics	Categories	EFA sample (n=281)		CFA sample (n=247)	
		Frequency (n)	Percent (%)	Frequency (n)	Percent (%)
Gender	Female	173	61.6	149	60.3
	Male	108	38.4	98	39.7
Education level	Bachelor's degree	211	75.1	183	74.1
	Master's degree	70	24.9	64	25.9
Work experience	<15 years	66	23.5	54	21.9
	15–20 years	58	20.6	45	18.2
	21–25 years	67	23.8	54	21.9
	26–30 years	45	16.0	60	24.3
	>30 years	45	16.0	34	13.8
School type	National standard	110	39.1	96	38.9
	Non-standard	171	60.9	151	61.1
Location	Urban	164	58.4	146	59.1
	Rural	117	41.6	101	40.9

The four PDCA dimensions were comprehensively covered by an initial pool of 28 items. Intentional overgeneration was implemented to ensure content representativeness and empirical refinement. In a pilot study that involved 52 in-service primary school teachers, each item was assessed. The pilot phase was designed to evaluate the distributional properties, preliminary internal consistency, item–total correlations, and item clarity, as well as potential ceiling effects. If an item met one or more of the following criteria, it was eliminated: i) an item–total correlation below 0.30, which indicates a weak contribution to the construct; ii) severe ceiling effects, defined as 80% or more responses at the maximum scale value; or iii) ambiguity reported by more than 15% of participants. Eight items were eliminated in accordance with these criteria. The final instrument is composed of 20 items, which are evenly distributed across the four PDCA constructs, with five items per dimension, as shown in Table 2. This balanced structure maintains factorial symmetry and facilitates subsequent factor-analytic validation procedures. The five-point Likert scale, which ranges from 1 (strongly disagree) to 5 (strongly agree), is employed to evaluate all PSTQS items. To achieve a satisfactory level of variance for psychometric analysis, this response format was chosen to balance respondent burden and sensitivity. The reliability testing and factor analysis, particularly EFA and CFA, are facilitated by the symmetrical scaling structure. The factorial balance is guaranteed by the equitable distribution of five items across each of the four PDCA dimensions, which also bolsters the specification of a four-factor measurement model. This structural symmetry improves the instrument's interpretability and enhances its suitability for subsequent structural modeling and validation procedures.

A two-round Delphi process was implemented to improve methodological transparency and enhance content validity. The expert panel was composed of four specialists who have over 15 years of professional experience, hold doctoral degrees in educational management, and are directly involved in the implementation of curricula or quality assurance processes. Experts assessed each item against a four-point relevance scale: 1 (not relevant), 2 (somewhat relevant), 3 (quite relevant), and 4 (highly relevant). The calculation of the content validity index (CVI) is recommended to be conducted using a four-point scale, as it eliminates a neutral midpoint [28]. In addition to quantitative ratings, experts provided qualitative feedback

on behavior specificity, conceptual alignment, and wording clarity. Items that achieved an item-level (I-CVI) of 0.78 or higher were provisionally retained during the initial Delphi round. An I-CVI of 0.75 indicates that at least three experts have rated the item as either 3 or 4. This result is due to the panel size of four experts. Items that scored between 0.60 and 0.75 were revised and resubmitted for evaluation, while those that scored below 0.60 would have been eliminated; however, no items met this threshold for removal. The primary objective of the revisions was to enhance alignment with PDCA constructs and clarify behavioral phrasing. Revised items were reassessed in the subsequent round. I-CVI values of 0.80 or higher were achieved by all 20 items. The scale-level CVI, as determined by the average method (S-CVI/Ave), was 0.92. The adequacy of construct representativeness and clarity is confirmed by the fact that this value surpasses the recommended benchmark of 0.90 for strong overall content validity [28]. The Delphi process collectively offers systematic evidence that the PSTQS exhibits robust content validity and theoretical coherence.

Table 2. Measurement items for the PSTQS

PDCA dimensions	Item codes	Item content
Plan	p1	I identify lesson objectives that are aligned with the required learning outcomes.
	p2	I design teaching content that closely follows the curriculum.
	p3	I select teaching methods appropriate to the objectives and content.
	p4	I plan learning activities suited to primary school students.
	p5	I prepare teaching materials and supportive resources adequately before class.
Do	d1	I conduct warm-up activities to engage students' interest in learning.
	d2	I organize knowledge-building activities according to the lesson design.
	d3	I guide students in practice activities to consolidate their knowledge.
	d4	I create opportunities for students to apply knowledge in real-life situations.
	d5	I manage the classroom with a positive, student-centered approach.
Check	c1	I observe and pose questions to assess students' participation.
	c2	I use exercises or learning products to evaluate learning outcomes.
	c3	I provide students with specific and timely feedback.
	c4	I assess each student's progress compared with their own previous performance.
	c5	I adjust teaching activities based on the evaluation results.
Act	a1	I analyze the strengths and limitations of the lesson after implementation.
	a2	I exchange ideas with colleagues to draw teaching experience.
	a3	I update new teaching methods and techniques to improve quality.
	a4	I revise lesson plans for subsequent teaching.
	a5	I participate in professional development and training to enhance teaching competence.

To improve transparency and construct traceability, Table 3 explicitly articulates the relationship between the PSTQS items, GEP-derived pedagogical expectations, and PDCA theory. This mapping elucidates the manner in which each dimension of the PDCA cycle was translated into measurable teaching behaviors that were consistent with Vietnam's competency-based reform framework. The table illustrates that the scale is not a generic adaptation of the PDCA nor a simple restatement of curriculum requirements, but rather an integrated operational model that is rooted in both traditions. Such an outcome is achieved by connecting abstract quality management principles to concrete instructional practices. The PDCA function of defining objectives and designing processes is correlated with the Plan dimension. In the context of the classroom, this aspect entails the alignment of lesson objectives with the necessary learning outcomes, the maintenance of curriculum coherence, the selection of suitable teaching methods, the planning of age-appropriate learning activities, and the preparation of instructional materials in advance. Both GEP policy mandates and prior scholarship on instructional design and quality planning substantiate these behaviors, as reflected in items p1–p5.

The Do dimension reflects the execution of instructional plans. This part involves organizing tasks that help students learn, creating activities to keep them engaged, guiding their practice, helping them apply what they learn in real life, and maintaining a positive classroom environment focused on students. Items d1–d5, grounded in pedagogy, classroom management, and student-centered learning research, capture these classroom-level enactments of planned instruction. Monitoring and evaluation of outcomes are represented by the Check dimension. This aspect is consistent with formative assessment practices within the GEP framework, including the observation of participation, the evaluation of learning products, the provision of timely feedback, the assessment of individual progress in relation to prior performance, and the modification of instruction based on evaluation results. The literature on quality assurance and assessment for learning substantiates these monitoring and feedback mechanisms, operationalized in items c1–c5. Lastly, the Act dimension represents the combination of corrective action and reflective improvement. This notion encompasses post-lesson analysis, collegial exchange of experience, updating instructional methods, revising lesson plans, and participating in professional development activities in the context of teaching. Items a1–a5

capture these behaviors as iterative improvement processes essential to both PDCA theory and teacher professionalism. This structured mapping illustrates the conceptual coherence between quality management theory and classroom-level instructional practice. It verifies that each item is theoretically anchored, contextually grounded, and systematically aligned with a single PDCA stage, thereby enhancing the theoretical integrity and construct validity.

Table 3. Theoretical mapping of PDCA dimensions to teaching behaviors and items

PDCA dimensions	Core PDCA function	Operational teaching behaviors (GEP-aligned)	Item codes	Key sources
Plan	Define objectives and design processes	Align lesson objectives with required outcomes; curriculum-based content design; appropriate method selection; age-suited activity planning; material preparation	p1–p5	[1], [6], [9], [12], [13], [16], [20], [29], [30]
Do	Execute instructional plan	Conduct engagement activities; organize knowledge-building tasks; guide practice; promote real-life application; manage student-centered classroom	d1–d5	[4]–[6], [10], [16], [19], [29], [31], [32]
Check	Monitor and assess outcomes	Observe participation; assess via tasks/products; provide timely feedback; evaluate individual progress; adjust teaching based on assessment	c1–c5	[11], [13], [14], [33]–[36]
Act	Reflect and implement improvement	Post-lesson analysis; collegial exchange; update methods; revise lesson plans; engage in professional development	a1–a5	[4], [8], [10]–[13], [29], [37], [38]

To mitigate the potential impact of common method bias (CMB), both procedural and statistical remedies were implemented, as the study was dependent on self-reported survey data collected through a single instrument. Survey design included numerous procedural safeguards. Initially, the order of the items was randomized to mitigate consistency artifacts and pattern responding. Secondly, to reduce the impact of social desirability and the apprehension associated with evaluation, all items were composed using neutral and behavior-focused language. Third, respondents were guaranteed anonymity and confidentiality, which diminished the probability of inflated self-assessments that were motivated by perceived evaluation risk. To evaluate the existence of common method variance, Harman's single-factor test was implemented. No single factor accounted for more than 30% of the total variance, as indicated by the unrotated EFA. This finding implies that construct validity was not at risk of being significantly compromised by CMB. The confidence in the integrity of the measurement model and the credibility of subsequent analyses is enhanced by the combined effect of the procedural and statistical controls.

3.4. Data analysis

Data the psychometric properties and factor structure of the PSTQS were validated using Jamovi (Version 2.6) for data analysis. To facilitate cross-validation, the validation procedure was divided into two sequential stages: EFA and CFA, which were each conducted on two independent subsamples. The initial subsample (n=281) was analyzed during the EFA phase to determine the structure of the fundamental factors. Two diagnostic tests were implemented to assess the dataset's suitability for factor analysis prior to extraction. These included Bartlett's test of sphericity, which required a statistically significant result ($p<0.05$), and the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy, which had a minimum acceptable value of 0.50. After the data adequacy was verified, latent factors were extracted using principal axis factoring (PAF) with Promax rotation. Items with factor loadings below 0.50 were considered for removal to ensure construct clarity and factor purity, while factors with eigenvalues greater than 1.0 were retained. The second subsample (n=247) was employed to test the four-factor model, which was derived from EFA, during the CFA stage. The model fit was evaluated using a variety of established fit indices, with the following cutoff criteria: Chi-square/df ratio (χ^2/df)≤3.00, comparative fit index (CFI) ≥0.90, Tucker-Lewis index (TLI) ≥0.90, root mean square error of approximation (RMSEA) ≤0.08, and standardized root mean square residual (SRMR) ≤0.08. The scale's reliability and validity were assessed subsequent to model confirmation. Composite reliability (CR) was employed to evaluate internal consistency reliability, with CR values of 0.70 or higher being deemed acceptable. With a target threshold of average variance extracted (AVE) ≥0.50, convergent validity was assessed using the AVE. Discriminant validity was evaluated using the Heterotrait-Monotrait ratio (HTMT), with construct separation values below 0.85 indicating satisfactory separation. The robustness, reliability, and construct validity of the PSTQS across independent samples were guaranteed by this multi-step analytic strategy.

4. RESULTS

4.1. EFA results

In order to investigate the latent structure of the 20-item PSTQS, an EFA was implemented on the initial subsample. The dataset's suitability for factor analysis was verified through preliminary diagnostics. The KMO value of 0.79 suggested that the items shared a sufficient common variance for factor extraction, indicating meritorious sampling adequacy. The null hypothesis that the correlation matrix was an identity matrix was rejected by Bartlett's test of sphericity, which was statistically significant ($\chi^2(190)=1782$, $p<0.001$). The combination of these indicators served as a compelling empirical rationale for the continuation of factor analysis. In accordance with the theoretical objective of identifying the underlying dimensions of teaching quality, PAF was chosen to extract latent constructs based on shared variance. Promax rotation was implemented to accommodate correlated factors as a result of the conceptual interdependence of PDCA components. The analysis produced a four-factor solution that was in accordance with the hypothesized PDCA framework. The four factors collectively accounted for 45.3% of the total variance, and each factor had an eigenvalue greater than 1.0. This proportion of explained variance is satisfactory and consistent with prior educational measurement studies for a multidimensional construct such as teaching quality, which is characterized by contextual variability and behavioral diversity.

A theoretically coherent and clean structure was revealed by the rotated factor matrix. The factorial purity of all items is supported by the fact that they loaded strongly onto their intended factors (0.53–0.77), with no substantial cross-loadings, as seen in Table 4. Factor 1 (Do) represented the enactment of instructional practice, Factor 2 (Act) reflected reflective and developmental behaviors, Factor 3 (Plan) represented preparatory instructional design, and Factor 4 (Check) encompassed assessment-related practices. The alignment between empirical grouping and theoretical expectation provides strong preliminary evidence for construct validity. Further, descriptive statistics provide valuable insights. Teachers reported frequent engagement in instructional delivery and assessment-related activities, as evidenced by the comparatively higher mean scores of items within the Do and Check dimensions. This pattern may be indicative of the practical realities of classroom teaching, in which instructional execution and student evaluation are immediate, visible, and institutionally monitored responsibilities. On the other hand, the relatively moderate means observed in the Plan and Act dimensions indicate that reflective refinement and long-term professional improvement activities may occur in a less systematic manner. This distribution is theoretically informative in that it implies that teachers exhibit proficiency in the operational phases of the PDCA cycle (Do and Check), but the cyclical improvement component (Act) may be less institutionalized, indicating a need for professional development. The inter-factor correlations varied from –0.04 to 0.30. The modest magnitude of these correlations serves as confirmation that the four constructs are empirically distinct while remaining moderately related. This pattern is consistent with the cyclical logic of the PDCA: each phase functions independently but sequentially contributes to a continuous improvement process. The absence of high inter-factor correlations (>0.70) further suggests that the dimensions are not redundant, thereby reinforcing the multidimensional nature of teaching quality as conceptualized in this study. In general, the EFA results substantiate the hypothesized four-factor structure and provide helpful information about the practical manifestations of the various phases of the instructional improvement cycle.

Table 4. Descriptive statistics and factor loadings for the EFA solution (n=281)

PDCA dimensions	Item codes	Mean	SD	Factor loading
Do	d1	4.16	0.79	0.72
	d2	4.20	0.73	0.74
	d3	4.22	0.75	0.59
	d4	4.19	0.74	0.77
	d5	4.18	0.74	0.53
Act	a1	3.96	0.76	0.56
	a2	4.03	0.81	0.62
	a3	3.75	0.82	0.70
	a4	3.88	0.79	0.77
	a5	3.97	0.83	0.68
Plan	p1	3.49	0.80	0.65
	p2	3.71	0.88	0.66
	p3	3.54	0.78	0.70
	p4	3.51	0.76	0.69
	p5	3.66	0.83	0.59
Check	c1	4.15	0.69	0.59
	c2	4.16	0.77	0.56
	c3	4.20	0.74	0.65
	c4	4.13	0.73	0.68
	c5	4.15	0.76	0.62

4.2. Confirmatory factor analysis results

To confirm the four-factor structure identified through EFA, a CFA was conducted on the second subsample (n=247) using maximum likelihood estimation. The hypothesized measurement model, which is based on the PDCA framework and consists of four correlated latent constructs (PDCA), exhibited an exceptional overall model fit, as shown in Table 5. A strong relative model fit was indicated by the Chi-square to degrees of freedom ratio (χ^2/df=1.49), which was significantly lower than the recommended threshold of 3.0. The CFI=0.951 and TLI=0.943 both exceeded the 0.90 benchmark for acceptable fit and approached the more stringent 0.95 criterion, further supporting the model's adequacy. The absolute fit indices were also robust: the RMSEA=0.04, 90% CI [0.03, 0.05]) was below the 0.06 cutoff for close fit, and the SRMR=0.05 remained comfortably under the 0.08 threshold. Collectively, these indices suggest that the four-factor PDCA model is an exemplary fit for the observed data. The statistical significance of all standardized factor loadings was significant ($p<0.001$), with a range of 0.54 to 0.79, as shown in Figure 2. The magnitude and consistency of these loadings serve as confirmation that each item makes a substantial contribution to its intended latent construct. The structural integrity of the measurement model was further substantiated by the absence of cross-loadings or specification errors.

Table 5. Goodness-of-fit indices for the CFA model (n=247)

Fit index	Recommended value	Model value	Interpretation
χ2/df	<3	1.49	Excellent fit
CFI	>0.90	0.95	Excellent fit
TLI	>0.90	0.94	Good fit
RMSEA	<0.06	0.04	Excellent fit
90% CI for RMSEA	Lower<0.06, Upper<0.08	[0.03, 0.05]	Excellent fit
SRMR	<0.08	0.05	Excellent fit

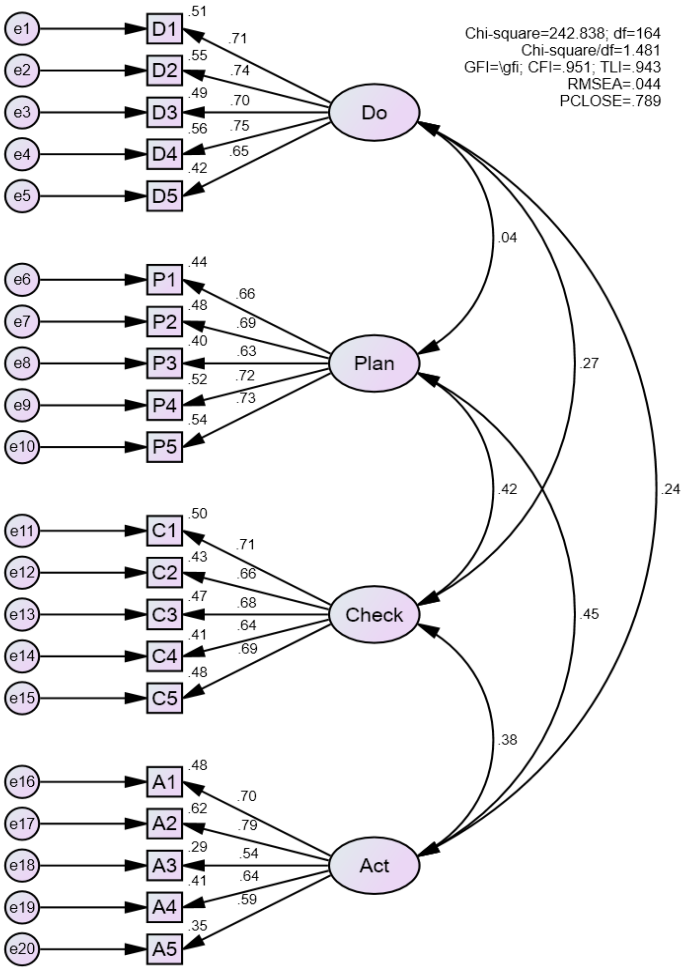

Figure 2. The confirmed four–factor model of the PSTQS with standardized estimates

The measurement model's results are illustrated in Table 6. The CR values were robust across all constructs: Plan (0.82), Do (0.84), Check (0.81), and Act (0.79), all of which exceeded the recommended threshold of 0.70. These findings suggest that the latent construct reliability and internal consistency are satisfactory. The AVE for the Do construct satisfies the recommended 0.50 criterion with regard to convergent validity (AVE=0.51). Nevertheless, the AVE values for Plan (0.48), Check (0.46), and Act (0.43) were marginally lower than the conventional 0.50 threshold. Methodological literature acknowledges that AVE values marginally below this cutoff may be acceptable when CR exceeds 0.70 and factor loadings are substantively strong, although AVE≥0.50 is frequently the recommended threshold [28]. The present study has shown that all three constructs have CR values that are significantly higher than 0.70, and the majority of standardized loadings exceed 0.60. This suggests that the indicators have an adequate shared variance. It is crucial to note that the slightly lower AVE values are likely due to the multidimensional and behaviorally diverse nature of classroom practice. "Act" (teaching improvement) is a construct that inherently encompasses heterogeneous but theoretically coherent behaviors, such as professional development, reflection, and collaboration. While maintaining theoretical completeness, this conceptual breadth may modestly reduce shared variance. The AVE would have been marginally increased by removing lower-loading items (e.g., a3=0.54); however, such an action would have resulted in a reduction in construct coverage and a loss of theoretical fidelity to the PDCA improvement cycle. To assess this trade-off, additional diagnostic tests were implemented. The simulations of item removal demonstrated that the deletion of lower-loading items resulted in trivial AVE increases (≤0.03) and weakened content representativeness. Consequently, item retention was theoretically justified to maintain the integrity of the construct and ensure that it is consistent with the PDCA framework.

Table 6. Results of the measurement model: loadings, reliability, and convergent validity (n=247)

Construct	Item	β	S.E.	CR	AVE
Plan	p1	0.66	-	0.82	0.48
	p2	0.69	0.11		
	p3	0.63	0.10		
	p4	0.72	0.10		
	p5	0.73	0.11		
Do	d1	0.71	-	0.84	0.51
	d2	0.74	0.09		
	d3	0.70	0.09		
	d4	0.74	0.10		
	d5	0.65	0.10		
Check	c1	0.71	-	0.81	0.46
	c2	0.65	0.11		
	c3	0.68	0.10		
	c4	0.64	0.10		
	c5	0.69	0.10		
Act	a1	0.69	-	0.79	0.43
	a2	0.79	0.12		
	a3	0.54	0.11		
	a4	0.63	0.11		
	a5	0.58	0.10		

Note. All standardized factor loadings (β) are significant at p<0.001

Additional interpretive insight is provided by discriminant validity analysis, as seen in Table 7. Each PDCA phase is empirically distinct, as confirmed by the Fornell–Larcker criterion and HTMT ratios. The moderate correlation between Plan and Act shows their conceptual proximity. Planning informs subsequent revisions, and reflective improvement feeds forward into subsequent planning cycles. This moderate correlation empirically substantiates the cyclical and iterative characteristics of the PDCA framework, rather than suggesting redundancy. The CFA findings, when considered collectively, confirm that teaching quality is most effectively represented as four distinct yet interrelated phases, as conceptualized through a continuous improvement lens. The reliability estimates, validity evidence, and model fit indices collectively demonstrate that the PSTQS captures the structural differentiation and dynamic interconnection that are inherent in the PDCA framework. The results substantively indicate that instructional quality in primary education operates as a structured cycle, with strong implementation and evaluation, beyond statistical adequacy. Reflective refinement may necessitate additional institutional strengthening.

Table 7. Discriminant validity: Fornell-Larcker criterion and HTMT ratios

Construct	Plan	Do	Check	Act
Plan	0.69			
Do	0.07	0.71		
Check	0.38	0.23	0.67	
Act	0.40	0.19	0.35	0.66

Note: Diagonal values are the square root of the AVE. Off-diagonal values are inter-factor correlations (below diagonal) and HTMT ratios (above diagonal).

5. DISCUSSION

The primary objective of this study was to establish and verify the PSTQS as a PDCA-based framework for continuous instructional improvement in Vietnamese primary education. The discussion emphasizes the theoretical and practical implications of these results for the implementation of educational quality management and reform, rather than merely reiterating statistical findings. The current investigation makes three substantial contributions to the existing body of literature. Initially, it facilitates the conceptual integration of classroom-level pedagogy and TQM. Prior studies have advocated for the PDCA cycle as an effective framework for improving educational quality [8]–[11]; however, its empirical application in daily teaching practice has been limited. The PSTQS bridges the gap between micro-level pedagogical processes and organizational quality management theory by translating the PDCA into observable instructional behaviors. This bolsters the argument that continuous improvement models are not merely administrative tools; they can also significantly influence classroom practice. Secondly, the results serve to underscore the complexities of teaching quality. The empirical distinction between PDCA implies that instructional quality is not a single global competence but rather a cyclical process that involves preparation, enactment, evaluation, and refinement. The moderate inter-factor correlations further support the interpretation of teaching quality as a dynamic improvement loop rather than a linear performance construct. This is consistent with the iterative improvement logic that is emphasized in the PDCA scholarship [8], [10] and bolsters theoretical arguments that quality in education is maintained through structured reflection and feedback mechanisms. Third, the investigation addresses a methodological gap in the Vietnamese context. The PDCA was typically discussed conceptually or applied within localized school improvement initiatives in previous national studies [12], [13]. Nevertheless, there was a dearth of psychometrically validated and standardized instruments that were appropriate for system-wide monitoring. This study addresses the requests of Rodríguez–Mantilla *et al.* [15] and Wafudu *et al.* [16] for more robust measurement tools to bolster the IQAS by creating a rigorously validated scale. In doing so, it places Vietnam's reform endeavors within the broader international discourse on evidence-based quality governance.

The 2018 GEP prioritizes reflective professionalism, formative assessment, and competency-based instruction. The PSTQS provides a concrete means for assessing the practical application. It is important to point out that the EFA descriptive results show a pattern where “Do” and “Check” are more popular than “Act.” This conclusion suggests that teachers may be more confident in how they teach and grade students than in how they improve their own work through systematic reflection. of these reform principles. It is crucial to note that the pattern observed in the EFA descriptive results—a higher endorsement of “Do” and “Check” in comparison to “Act”—indicates that teachers may be more assured in the delivery of instruction and student assessment than in the process of systematic reflective improvement. This finding has theoretical implications: the implementation of reform may initially enhance visible instructional behaviors before fully institutionalizing reflective and collaborative professional cultures. This observation is consistent with international research that indicates that the implementation of continuous improvement frameworks necessitates sustained cultural change rather than procedural compliance [3]. Therefore, the PSTQS is capable of indicating developmental imbalances within the PDCA cycle, which can assist policymakers in determining which phases of the improvement process need support. It does not merely assess current practice.

The PSTQS serves as a structured reflective tool at the teacher level. In contrast to generic evaluation checklists, its PDCA alignment motivates educators to view instruction as an iterative improvement process. This aspect is consistent with the recommendations of Bogren *et al.* [14], who underscore the importance of self-regulated, evidence-informed teaching in the maintenance of professional development. Aggregated PSTQS data at the school leadership level can inform targeted professional development strategies. For example, systematic problems in the "Check" dimension may show that teachers need training in formative assessment practices. This approach is in line with the IQAS methods that Kaplani and Zafiropoulos [18] promote. The scale can connect national reform goals with clear teaching measures because it is based on both PDCA theory and GEP requirements. As such, the instrument advocates for data-driven school improvement planning rather than episodic intervention. The MOET may employ the PSTQS as a monitoring tool to evaluate the fidelity of the GEP implementation at the policy level. The scale

is able to align national reform objectives with measurable instructional indicators due to its foundation in both PDCA theory and GEP mandates. This requirement is in accordance with international quality assurance recommendations that emphasize accountability systems that are substantiated by dependable metrics [5], [17]. Positioning the PSTQS as a developmental instrument, rather than a high-stakes evaluation tool, is crucial. International quality management research warns that punitive perception renders improvement-oriented instruments ineffective [9], [11]. Consequently, sustainable implementation will necessitate clear communication and supportive policy framing.

In spite of its advantages, it is imperative to recognize numerous constraints. The PSTQS was initially validated in Vietnamese primary schools that were operating under the 2018 GEP. Contextual adaptation may be required prior to the implementation of the PDCA in other educational systems, despite its global recognition. Cultural expectations regarding autonomy, collegial collaboration, and reflection may influence response patterns and necessitate recalibration. Secondly, the PSTQS is a self-report instrument that captures perceived instructional behaviors rather than externally observed practices. Although statistical and procedural remedies were implemented to mitigate CMB, future research could incorporate student outcome data or classroom observation to enhance criterion-related validity. Third, institutional readiness may influence the success of implementation. Schools that lack administrative support or have restricted professional learning communities may encounter difficulties in operationalizing the "Act" phase of continuous improvement. The PSTQS captures structured instructional processes, but it does not directly measure structural constraints. Therefore, it is crucial to consider the wider systemic conditions when interpreting the results. This investigation transcends the mere verification of a statistical model and illustrates how the PDCA framework can be operationalized as a context-sensitive, coherent, and measurable representation of teaching quality. The PSTQS makes a practical and theoretical contribution to the field of educational quality assurance by combining quality management theory with classroom practice and national reform mandates. Its value is not solely derived from its precision in measurement but also from its ability to conceptualize teaching as a disciplined cycle of planning, implementation, evaluation, and improvement—an orientation that is essential for the sustainable transformation of education.

6. CONCLUSION

The aim of this study was to assess the quality of primary school teaching in alignment with Vietnam's 2018 GEP using a scientifically validated instrument. The research provides a theory-driven, robust measure of instructional practices through the development and validation of the PSTQS, a 20-item instrument that is rooted in the PDCA framework. The two-phase design, which included EFA and CFA with a large, independent sample of 528 in-service teachers, confirmed the scale's four-factor structure. The design also exhibited strong convergent and discriminant validity, as well as high internal consistency. The study's primary contribution is the transformation of the PDCA cycle from a theoretical concept to a practical, psychometrically sound instrument. The PSTQS offers a concrete, actionable tool to guide and support educational improvement in the face of curriculum reforms that require more than policy rhetoric. It provides policymakers with evidence to monitor instructional quality under the GEP, empowers teachers to engage in structured self-reflection, and enables school leaders to design data-informed professional development programs. By doing so, the PSTQS establishes a clear path to enhance the effectiveness of teaching and encourage the development of student competency, thereby bridging the gap between policy and practice.

Nevertheless, there are several methodological constraints that necessitate more explicit consideration, particularly in the context of sample representativeness. Initially, the study implemented convenience sampling through the distribution of online Google Forms. Although this method facilitated the efficient collection of data and widespread participation, it also introduced the possibility of self-selection bias. Teachers who elected to respond may have been more optimistic about quality improvement initiatives, professionally engaged, or technologically confident. As a result, the sample may overrepresent educators who are already reflective or motivated, which could potentially inflate reported engagement across PDCA dimensions—particularly in areas such as planning and assessment. Additionally, the exclusion of teachers with limited digital access or lower technological literacy due to online administration may have impacted the dataset. Consequently, the diversity of instructional contexts may have been reduced. The demographic composition of the sample also suggests a possible bias in favor of experienced and urban-based educators. Factor loadings and mean scores may be influenced by the more stable instructional routines and greater familiarity with curriculum reform expectations that highly experienced educators may report. Experience can make professional judgment better, but it can also hide the differences between new teachers who are still learning how to use PDCA-aligned skills. Consequently, the factorial structure that was confirmed in this study may be indicative of a professional cohort that is relatively mature, rather than the complete spectrum

of teacher developmental stages. Third, data were exclusively collected from Ho Chi Minh City, a metropolitan area that is distinguished by comparatively robust institutional support, greater access to professional development, and superior infrastructure in comparison to numerous rural or remote regions of Vietnam. In urban settings, schools frequently receive advantages from more consistent policy implementation, stronger leadership capabilities, and more extensive peer collaboration networks. In contrast, rural schools may encounter resource constraints, larger multigrade classrooms, and limited professional learning opportunities, all of which could potentially impact the implementation of PDCA cycles, particularly the "Act" dimension, which involves professional development and reflection. Consequently, it is advisable to exercise prudence when applying these findings to a national scale. Future research should address these concerns by ensuring balanced representation based on teacher experience and school type, incorporating rural and disadvantaged schools, and employing stratified random sampling across regions. This replication would enhance the PSTQS's national applicability and external validity.

FUNDING INFORMATION

This research received no specific financial support.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Nhat Thong Du	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Van Dat Tran	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

INFORMED CONSENT

Informed consent was obtained from all subjects involved in the study.

ETHICAL APPROVAL

The research related to human use has been complied with all the relevant national regulations and institutional policies in accordance with the tenets of the Helsinki Declaration and has been approved by the by the ethical review committee of Ho Chi Minh City University of Education.

DATA AVAILABILITY

The data that supports the findings of this study are available on request from the corresponding author, [VDT]. The data, which contains information that could compromise the privacy of research participants, is not publicly available due to certain restrictions.

REFERENCES

[1] Ministry of Education and Training, "CIRCULAR: Promulgating General Education Program No. 32/2018/TT-BGDĐT," 2018. [Online]. Available: https://dulieuphapluat.vn/van-ban/linh-vuc-khac-van-ban/circular-322018tt-bgddt-on-the-general-education-program-issued-by-the-minister-of-education-and-training-1011743.html?

[2] M. M. Asad and N. Suleman, "Impact of Technology-Supported Personalized Learning 5.0 on Instructional Quality: Insights from the Higher Education Institutions of Pakistan," *Quality Assurance in Education*, vol. 33, no. 3, pp. 445–461, May 2025, doi: 10.1108/QAE-10-2024-0200.

[3] H. M. Mcheka and M. A. Mislay, "Politics of School Evaluation during Quality Assurance Policy Implementation in Selected Secondary Schools in Dodoma, Tanzania," *Quality Assurance in Education*, vol. 33, no. 3, pp. 500–519, May 2025, doi: 10.1108/QAE-07-2024-0125.
[4] G. Qingyan, A. S. Azar, and A. Ahmad, "The Impact of Teacher Quality Management on Student Performance in the Education Sector: Literature Review," *World Journal of English Language*, vol. 13, no. 3, pp. 156–171, Mar. 2023, doi: 10.5430/wjel.v13n3p156.
[5] K. Parveen *et al.*, "The Contribution of Quality Management Practices to Student Performance: Mediated by School Culture," *Heliyon*, vol. 10, no. 15, p. e34892, Aug. 2024, doi: 10.1016/j.heliyon.2024.e34892.
[6] E. Sfakianaki, A. Kakouris, and C. Siontorou, "Critical Success Factors for Total Quality Management in Primary and Secondary Education," *International Journal of Services and Operations Management*, vol. 40, no. 4, pp. 564–595, 2021, doi: 10.1504/IJSOM.2021.120061.
[7] C. W. Huan and N. B. Mohamad Nasri, "Teacher Teaching Practices Based on the PDCA Model: A Systematic Literature Review," *International Journal of Academic Research in Progressive Education and Development*, vol. 11, no. 3, Aug. 2022, doi: 10.6007/ijarped/v11-i3/14126.
[8] Q. Yan, L. Li, M. Chen, and X. Gu, "Research on the Sustainable Improvement Mechanism of the Chemical Engineering and Technology Major Based on the Concepts of Outcome-Based Education–Plan-Do-Check-Act (OBE–PDCA) in Engineering Education," *Sustainability (Switzerland)*, vol. 17, no. 9, p. 4051, Apr. 2025, doi: 10.3390/su17094051.
[9] A. Marini and L. Zakiah, "Comprehensive Factors Influencing the Success of Internal Quality Assurance Systems in Elementary Schools: Resources, Leadership, and Community Support," *Jurnal Pendidikan Administrasi Perkantoran*, vol. 13, no. 1, pp. 440–450, 2025, doi: 10.26740/jpap.v13n1.p440-450.
[10] M. Sri, J. I. S. P. Jenny, Suharno, and G. Rivan, "The Quality Management of Education in Elementary Schools in Improving Teachers' Digital Literacy in the Era of Online Learning," *Pegem Journal of Education and Instruction*, vol. 14, no. 1, pp. 29–37, Aug. 2023, doi: 10.47750/pegegog.14.01.04.
[11] C. J. La R. Longobardi, L. C. P. Rivera, L. V. A. Trujillo, C. F. A. Chávez, D. R. C. Armas, and S. F. F. Piñas, "Accreditation and the Deming Cycle in the Faculties of the National University of Education-2022," *Journal of Namibian Studies: History Politics Culture*, vol. 33, May 2023, doi: 10.59670/jns.v33i.1203.
[12] H. T. Hai and C. T. T. Xuan, "How to Enhance the Managerial Staff Capacity in Southeast Vietnam's Primary Schools for Educational Reform?" *Journal of Education and Learning*, vol. 15, no. 1, p. 38, Jul. 2025, doi: 10.5539/jel.v15n1p38.
[13] N. T. Viet Ha, L. H. Hoa, H. T. Trang, P. Le Thuy, and D. T. N. Huy, "Several Techniques and Methods for Improving English Teaching Quality for Students Based on Learner-Centered Approach," *International Journal of Language, Literature and Culture*, vol. 2, no. 5, pp. 12–17, 2022, doi: 10.22161/ijllc.2.5.3.
[14] M. Bogren, P. Jha, B. Sharma, and K. Erlandsson, "Development of a Quality Assurance Assessment Tool to Meet Accreditation Standards for Midwifery Education: A Delphi Study," *Women and Birth*, vol. 37, no. 5, p. 101660, Sep. 2024, doi: 10.1016/j.wombi.2024.101660.
[15] J. M. Rodríguez-Mantilla, M. J. Fernández-Díaz, and V. León Carrascosa, "Validation of a Questionnaire to Evaluate the Impact of ISO 9001 Standards in Schools with a Confirmatory Factor Analysis," *Studies in Educational Evaluation*, vol. 62, pp. 37–48, Sep. 2019, doi: 10.1016/j.stueduc.2019.03.013.
[16] S. J. Wafudu, Y. Bin Kamin, and D. Marcel, "Validity and Reliability of a Questionnaire Developed to Explore Quality Assurance Components for Teaching and Learning in Vocational and Technical Education," *Humanities and Social Sciences Communications*, vol. 9, no. 1, p. 303, Sep. 2022, doi: 10.1057/s41599-022-01306-1.
[17] E. A. Mtitu, "Exploring Factors for the Introduction and Implementation of Quality Assurance Systems in Selected Higher Education Institutions in Tanzania," *Cogent Education*, vol. 12, no. 1, Dec. 2025, doi: 10.1080/2331186X.2025.2455767.
[18] G. Kaplani and K. Zafiropoulos, "A Model of Quality Assurance in Primary Education Management. The Case of Greece," *Cogent Business and Management*, vol. 9, no. 1, Dec. 2022, doi: 10.1080/23311975.2021.2016555.
[19] J. Juharyanto, I. Arifin, S. Sultoni, M. A. Adha, and M. I. Qureshi, "Antecedents of Primary School Quality: The Case of Remote Areas Schools in Indonesia," *SAGE Open*, vol. 13, no. 1, Jan. 2023, doi: 10.1177/21582440221144971.
[20] E. Sucuoğlu and G. Erdem, "Effects of Sustainable Strategic Planning Applications in Primary Schools on the Effectiveness of Total Quality Management Practices," *Sustainability (Switzerland)*, vol. 13, no. 18, p. 9998, Sep. 2021, doi: 10.3390/su13189998.
[21] H. T. Huang and F. C. I. Chang, "Applying PDCA Cycle for Detecting Preschools' Teaching Quality," *ICIC Express Letters, Part B: Applications*, vol. 12, no. 6, pp. 489–497, 2021, doi: 10.24507/icicelb.12.06.489.
[22] D.-T. Do, C.-L. Le, and T.-V. Giang, "The correlation between internal quality assurance and the formation of quality culture in Vietnam higher education: a case study in Ho Chi Minh City," *European Journal of Educational Research*, vol. 9, no. 2, pp. 499–509, Apr. 2020, doi: 10.12973/eu-jer.9.2.499.
[23] H. Berbar, S. Lotfi, M. Essaoudi, and M. Talbi, "A Psychometric Study: The Validation of a School Quality Assessment Tool," *The Education and Science Journal*, vol. 24, no. 4, pp. 112–139, Apr. 2022, doi: 10.17853/1994-5639-2022-4-112-139.
[24] K. Nanthasudsawaeng, "Confirmative Component Analysis of Management Optimization of Educational Organizations," *Rajabhat Chiang Mai Research Journal*, vol. 24, no. 1, pp. 131–142, Jan. 2023, doi: 10.57260/rcmrj.2023.261086.
[25] M. S. Bajjou and A. Chafi, "Developing and Validating a New Conceptual Model for Successful Implementation of Lean Construction: SEM Analysis," *Engineering, Construction and Architectural Management*, vol. 32, no. 3, pp. 1581–1620, Mar. 2025, doi: 10.1108/ECAM-02-2023-0102.
[26] B. Saeed, R. Tasmin, A. Mahmood, and A. Hafeez, "Development of a Multi-Item Operational Excellence Scale: Exploratory and Confirmatory Factor Analysis," *TQM Journal*, vol. 34, no. 3, pp. 576–602, May 2022, doi: 10.1108/TQM-10-2020-0227.
[27] R. C. MacCallum, M. Roznowski, C. M. Mar, and J. V. Reith, "Alternative Strategies for Cross-Validation of Covariance Structure Models," *Multivariate Behavioral Research*, vol. 29, no. 1, pp. 1–32, Jan. 1994, doi: 10.1207/s15327906mbr2901_1.
[28] C. Fornell and D. F. Larcker, "Evaluating Structural Equation Models with Unobservable Variables and Measurement Error," *Journal of Marketing Research*, vol. 18, no. 1, pp. 39–50, 1921.
[29] Y. Wang and P. Chen, "Developing Quality Assurance System for College English Curriculum in Chinese Vocational Universities: A PDCA Cycle Model Approach," *European Journal of Education Studies*, vol. 12, no. 7, Jun. 2025, doi: 10.46827/ejes.v12i7.6088.
[30] K. Wu and N. Noichun, "The Quality Management Model of School-Based Curriculum in Primary Schools of Ganzi Tibetan Autonomous Prefecture," *Journal of Roi Kaensarn Academi*, vol. 9, no. 11, pp. 1807–1820, 2024.
[31] S. Koonkaew, Y. Yaboonthong, C. Puthaprasert, and P. Rupavijetra, "A Model of Academic Administration - Based on World Class Standard Schools in Secondary Educational Service Area," *Humanities, Arts and Social Sciences Studies*, vol. 19, no. 3, pp. 523–551, 2019, doi: 10.14456/hasss.2019.24.

[32] S. Romlah, A. F. Hamzah, and M. Makhzuniyah, "Qudwatuna as a Leadership Model in Education Quality Development," *Munaddhomah*, vol. 6, no. 3, pp. 404–419, 2025, doi: 10.31538/munaddhomah.v6i3.1871.
[33] X. Zhuang and M. Song, "Teaching Quality Evaluation and Improvement Based on Big Data Analysis," *Scalable Computing: Practice and Experience*, vol. 26, no. 1, pp. 220–230, 2025, doi: 10.12694/scpe.v26i1.3669.
[34] N. V. K. Jasti, V. Venkateswaran, S. Kota, and K. S. Sangwan, "A literature review on total quality management (models, frameworks, and tools and techniques) in higher education," *The TQM Journal*, vol. 34, no. 5, pp. 1298–1319, Nov. 2022, doi: 10.1108/TQM-04-2021-0113.
[35] J. Soria-García and Á. R. Martínez-Lorente, "Development and Validation of a Measure of the Quality Management Practices in Education," *Total Quality Management & Business Excellence*, vol. 25, no. 1–2, pp. 57–79, Feb. 2014, doi: 10.1080/14783363.2011.637790.
[36] N. Rahminawati and T. Supriyadi, "Implementing an Internal Quality Assurance System to Enhance Elementary School Education Quality," *International Journal of Learning, Teaching and Educational Research*, vol. 22, no. 4, pp. 414–433, Apr. 2023, doi: 10.26803/ijlter.22.4.23.
[37] T. Rostami Alileh, A. Rastgoo, Y. Namvar, and S. Sattari, "Validating of Applied Model for the Implementation of the School Management Excellence Program," *Iranian Journal of Educational Sociology*, vol. 5, no. 4, pp. 170–184, Dec. 2022, doi: 10.61186/ijes.5.4.170.
[38] R. Vajargah, K. Mousapour, and N. Tourani, "Designing and Validating a Model of Professional Learning Community for Secondary Technical and Vocational Schools," *Journal of School Administration*, vol. 11, no. 2, pp. 120–158, 2023.

BIOGRAPHIES OF AUTHORS

Nhat Thong Du obtained his Ph.D. in Educational Leadership from the National Taichung University of Education (NTCU), Taiwan, in 2016. He is currently a lecturer at the Faculty of Educational Sciences, Ho Chi Minh City University of Education, Vietnam. His main research directions are educational leadership and management, teacher professional development, competence-based education, teaching and learning methods, and instructional improvement. He can be contacted at email: nhatdt@hcmue.edu.vn; duthongnhat75@gmail.com.

Van Dat Tran holds a Ph.D. in Educational Leadership from La Trobe University, Australia, in 2013. He is a full professor at An Giang University, Vietnam National University Ho Chi Minh City, Vietnam. His research interests include educational leadership, organizational culture, school environment, teaching, and learning methods. He can be contacted at email: tvdat@agu.edu.vn; jptrandat@yahoo.com.