

Research Article

Cite this article: Sukkamart, A., Ployduangrat, J., Prongprommarat, J., & Pimdee, P. (2025). Development and Validation of a Competency Framework for Thai External Quality Assessment Supervisory Units. *Educational Process: International Journal*, 18, e2025454. <https://doi.org/10.22521/edupij.2025.18.454>

Received July 16, 2025

Accepted August 31, 2025

Keywords: Educational quality indicators; evaluation methods; program evaluation; school-based evaluation; Thailand

Author for correspondence:

Jaruwan Ployduangrat

✉ jaruwan.pl@kmitl.ac.th

✉ School of Industrial Education and Technology, King Mongkut's Institute of Technology Ladkrabang, Thailand.

**OPEN ACCESS**

© The Author(s), 2025. This is an Open Access article, distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution, and reproduction, provided the original article is properly cited.

Development and Validation of a Competency Framework for Thai External Quality Assessment Supervisory Units

Aukkapong Sukkamart , Jaruwan Ployduangrat , Jakkapong Prongprommarat , Paitoon Pimdee

Abstract

Background/purpose: Effective supervision is critical for credible educational quality assurance, yet a validated framework for External Quality Assessment Supervisory Units (EQASUs) remains absent in Thailand. This study aimed to develop and validate a competency framework of components and quality indicators to address this gap and enhance EQASU effectiveness.

Materials/methods: The research was conducted in two sequential stages, in which the first stage used 18 experts to analyze the framework's core components and indicators. Second, the model was empirically validated through a survey of 434 external quality assessors (EQAs) in 2023, selected via proportional stratified random sampling. Data were analyzed using Confirmatory Factor Analysis (CFA) to validate the construct and the Mann–Whitney U test to examine performance differences.

Results: The validation process confirmed a framework of four core components: organizational management (OM), personnel performance development (PPD), productivity and results (PAR), and cooperative networks (CN). All components demonstrated strong predictive power on EQASU effectiveness, with PPD emerging as the most crucial ($\beta = 0.94$), followed by OM ($\beta = 0.81$), CN ($\beta = 0.79$), and PAR ($\beta = 0.78$). A meta-evaluation of the framework's application confirmed the feasibility of its policy recommendations. Among the evaluation dimensions, utility was highest, followed by feasibility, accuracy, and propriety. Furthermore, high-performing EQASUs showed a significantly closer fit to the validated framework than the general group at the 0.01 level.

Conclusion: The verified EQASU framework addresses a critical need in the Thai educational landscape. It contributes to the literature on assessment supervision and offers a practical, evidence-based tool for strengthening external quality assurance in Thailand and similar international contexts.

1. Introduction

External Quality Assessments (EQAs) are essential in safeguarding and improving educational standards across countries, ensuring the accountability of schools and institutions to policy and society. Conducting such a national education quality auditing system, the External Quality Assessment Supervisory Unit (EQASU) is positioned with a vital role in operationalizing, monitoring, and nurturing assessors in combination with facilitating the reliability and validity of the assessors' performances (ONESQA, 2013, 2017, 2023; Pitiyanuwat, 2011). However, most education systems rarely establish systematic competency and quality assurance of EQASUs. However, in Thailand, these assessments have become critical to benchmarking school performance, guiding resource allocation, and fostering continuous institutional improvements (Tangwarasittichai et al., 2021).

Although EQAs have been implemented throughout Thailand for the last two decades, studies have found that supervisors' supervision appears to lack regularity, consistency, standardization, and evidence-based effectiveness indicators. Moreover, previous studies have reported that supervisors had unclear supervision roles, limited professional development, and a lack of collaborative cooperation between assessors and institutions (Mala, 2018; Saengpassa, 2014; Sriarunrasmee & Seekheio, 2024). Therefore, these weaknesses result in differences in the quality of the assessment and a lack of confidence in the results of the external assessment. These practical challenges in Thailand reflect a broader theoretical and empirical shortcoming.

While international researchers have emphasized the significant role of supervision and monitoring structures in upholding the trustworthiness of assessment (Ayeni, 2012; Kaiser et al., 2022; Pizzolato & Dierickx, 2023; Weldearegay, 2020), nonetheless, limited existing research has examined the design of supervisory unit frameworks and evaluated competency indicators with empirical validation, leaving a critical gap in the literature for the theorization and operationalization of EQASUs as a backbone of quality assurance.

Accordingly, the present study aimed to develop and validate a relatively holistic EQASU framework while identifying its important components and indicators for adequate quality assurance supervision (Chaiya & Ahmad, 2021; Henderson, 2025; Lee et al., 2023; Pitiyanuwat, 2011). This research provides theoretical and practical implications by triangulating expertise with empirical testing on many assessors. More specifically, this study tries to answer the following three research objectives (RO):

RO1: To identify core components of EQASU competency and supervision through expert consensus.

RO2: To empirically validate these components and their associated indicators using Confirmatory Factor Analysis (CFA).

RO3: To evaluate the practical utility of policy recommendations derived from the validated framework.

The study addresses these aims and contributes to the international literature on educational quality assurance. It provides policy-relevant insights for Thailand and other contexts seeking to strengthen supervisory mechanisms in external evaluation.

2. Literature Review

The study benefits ONESQA by improving institutional effectiveness by creating a validated framework explicitly developed for the Thai educational context. It also adds depth to international discourse on constructing and implementing supervision mechanisms for educational quality assurance. Given the widespread attempts to balance between accountability and development support in many education systems worldwide, insights from this study may help foster the

development of supervisory functions to improve quality assurance and evaluation systems around the world.

2.1. Organizational Management (OM)

Organizational management (OM) provides the structure through which EQASUs function effectively. An important quality of successful EQASUs is their adoption of a participative and systematic management structure that facilitates employee participation in and joint responsibility for OM.

Research findings detail how participatory decision-making can effectively increase motivation, job satisfaction, and institutional goals (Behraves et al., 2021; Galeazzo et al., 2021). In educational settings in particular, ensuring a culture of inclusion where all stakeholders are involved in management decisions is vital to developing sustainable quality assurance systems (Ma, 2023; Wang et al., 2022).

In Thailand, ONESQA has increasingly stressed the need to prioritize the contextual realities of the school and the university in their management strategies. By restructuring the EQA in its fourth round such that it entailed less paperwork and was more flexible, ONESQA tried to lessen the burden of bureaucracy from previous rounds (Sriarunrasmee & Seekheio, 2024; ONESQA, 2023). This transformation fits well in a broader international context wherein SAQA bodies across the globe are increasingly attentive towards responsive, decentralized, and participatory structures for management.

International evidence suggests EQA systems develop credibility when accountability is balanced with developmental support (El-Khawas, 2014; Harvey & Newton, 2007; Nadeem, 2023). Reports by UNESCO and the OECD stress that fragmented and/ or overly compliance-oriented systems undermine trust, particularly in emerging contexts (Henderson, 2025; Nguyen et al., 2020; United Nations Educational, Scientific, & Cultural Organization, 2016). This experience resonates with the challenges of the current Thai system and the importance of having a single framework for supervision.

2.2. Theoretical Frameworks for EQASU Organizational Design

The effectiveness of External Quality Assessment Supervisory Units (EQASUs) cannot be fully understood through empirical components alone; it requires a theoretical lens to analyze their underlying organizational design and legitimacy. This study draws primarily on Henry Mintzberg's configurational theory of organizational structure (Mintzberg, 1980) to provide this framework.

Mintzberg's theory proposes that organizations are structured around five core coordinating mechanisms and can be categorized into five ideal types. For EQASUs, two types are particularly relevant: the Professional Bureaucracy and the Machine Bureaucracy. A Professional Bureaucracy relies on the standardization of skills and the autonomy of highly trained experts (e.g., assessors making professional judgments). In contrast, a Machine Bureaucracy relies on standardizing work processes and centralized control (e.g., strict, uniform protocols for all evaluations) (Lunenburg, 2012). The tension between these models is central to the design of a QA system: it must balance the need for consistent, accountable processes (machine elements) with the need for adaptive, context-sensitive expert judgment (professional elements).

Furthermore, Institutional Theory (DiMaggio & Powell, 1983) helps explain the external pressures that shape EQASUs. Organizations often adopt structures and practices not solely for efficiency, but to gain legitimacy from their environment—a process known as isomorphism. This explains the global convergence of QA frameworks (e.g., towards models like EFQM) as EQASUs

respond to coercive (regulatory), mimetic (copying others), and normative (professional standards) pressures.

Together, these theories provide a robust foundation for analyzing the four key components of EQASU effectiveness discussed in the following sections. Organizational Management (OM) involves the strategic choice between mechanistic and organic structural configurations. Personnel Performance Development (PPD) is the mechanism for standardizing skills, essential for a Professional Bureaucracy. Productivity and Results (PAR) reflects the output control typical of a Machine Bureaucracy. Finally, Cooperative Networks (CN) can be understood as a response to institutional pressures for legitimacy and a mechanism for mitigating the isolation of professional autonomy. By applying this theoretical lens, this study moves beyond a descriptive list of components to an analytical examination of how EQASUs must be designed to navigate core organizational tensions.

2.3. Personnel Performance Development (PPD)

Personnel development is key to maintaining the competency of EQASU. Training, mentoring, and monitoring are continuous activities that enhance the practice of assessors in a manner that meets the institutional and national expectations. Evidence suggests a professional development uses formative assessment strategies and real-time feedback mechanisms. These, in turn, enhance the accuracy and comprehensiveness of assessors, rendering the assessment practice fairer and more balanced (Alam et al., 2022; Harris et al., 2017; Burgess & Phillips, 2021). Mentorship system (Roengsumran, 2015) and structured feedback loop mechanisms (Child & Ellis, 2021) ensure that assessors have enough opportunities to improve their skills and remain current with the ever-evolving evaluation framework changes.

Several national and regional projects have taken place in Thailand to professionalize EQASUs through systematic training and clear ethical and cultural standards, which continue to be underdeveloped elements of EQASUs practice in several contexts around the world (Asampong et al., 2023; Tangwarasittichai et al., 2021). Successful implementation of training in these areas ensures evidence of awareness and adherence to ONESQA's new and aligned standards, centering ethical and practical assessment.

Research in both Europe and Australia shows that much of the authority common QA outcomes generate as part of a credible accountability scheme depends on the intellectual, professional, and especially, the independent capacities of the oversight structure (Chikazinga, 2025; Dill, 2020; Kohoutek & Westerheijden, 2014; Newton, 2010). Without a strong supervisory unit that can provide professionalism and independence, institutional QA efforts risk degenerating into an exercise in compliance—a bureaucratic drag to be endured rather than a driver of improvement (Krooi et al., 2024; Stensaker, 2008). The credibility debate in QA situates EQASUs at the heart of generating meaningful QA outcomes in Thailand.

2.4. Productivity and Results (PAR)

The productivity of EQASUs can be measured by the speed with which they deliver evaluation reports and the extent to which the quality of such reports is satisfactory. Efficient EQASUs produce well-documented, actionable evaluation reports that help schools improve within strict timeframes (Preston et al., 2020). Delays and/or poorly written reports undermine the trustworthiness of QA systems (Narayanamma & Lalitha, 2016).

However, sometime the outcome is not what was expected (Brady & Bates, 2016; Kohoutek & Westerheijden, 2014), with an excellent case study of this coming from Germany, where Geisinger (1994) reported that the analysis of HE assessment in Germany highlighted a deplorable assessment

situation where the situation ended up decreasing public investment in the government's scientific and educational funds.

Therefore, internationally, the trend has been towards increasingly 'light-touch' evaluations, prioritizing evaluation for improvement and utility, at the expense of evaluation for administrative accountability. This trend has been particularly notable in Thailand, where the excessive documentation demanded during the third evaluation cycle was criticized for burdening teachers and school administrators (Mala, 2018; Saengpassa, 2014). The revised fourth round of ONESQA evaluations incorporated contextually responsive and strategic feedback, following international best practices (Pitiyanuwat et al., 2018).

Scholars argued that having supervisory competence requires more than the technicianship (technical expertise) of the evaluation process (Castanelli et al., 2022), as the process also involves organizational, interpersonal, and ethical competencies (Duarte & Vardasca, 2023; Schwarz & Westerheijden, 2007; Loukkola & Zhang, 2010). Therefore, these competences ensure that QA generates credibility but also fairness. The results of this study validate these arguments, as the four validated components of OM, PPD, PAR, and CN were compatible with these competencies in a holistic manner, reflecting the wider literature.

2.5. Cooperative Networks (CN)

Cooperative networks are crucial to the successful operation of EQASUs. Coordination among assessors, regulatory authorities, and universities ensures the consistency and transparency of the quality assurance process (Sriputtarin, 2008; Zheng et al., 2020). Communication mechanisms that enable the steady exchange of information and joint decision-making play an important role in contexts where assessments impact institutional prestige and access to funds (Sriarunrasmee & Seekheio, 2024; Weldearegay, 2020).

International comparisons suggest that countries with organic networks of assessors demonstrate higher fidelity and stakeholder satisfaction (Kaiser et al., 2022). In Thailand's new EQA model, these cooperative structures foster a more collegial and formative assessment process, breaking down the combative atmosphere created by some external evaluations (Valentine et al., 2021).

Evidence from research highlights the centrality of professional development in the maintenance of assessor competence and the credibility of QA processes (Brennan & Shah, 2000; Leiber et al., 2015; Stensaker & Leiber, 2015). The absence of structured opportunities for training increases the likelihood of inconsistencies in judgment and practice. Our findings' relevance to personnel performance development replicates this international consensus.

These four elements are grounded in internationally recognized quality assurance and supervisory frameworks. In detail, OM (Organizational Management) was derived from and correspond to the definition of governance and structural accountability included in specific models (e.g. EFQM, Baldrige); PPD (Personnel Performance Development) was in alignment with human capital and assessor professionalization accentuated in ONESQA's guidelines and the global literature on quality assurance, PAR (Productivity and Results) reflected performance and results-based accountability framework regularly employed in both educational and evaluation studies and CN (Cooperative Networks) was related to collaborative and stakeholder engagement dimension put forward in so many inter-organizational quality assurance research projects. This research validated continuity with the existing literature by foregrounding OM, PPD, PAR, and CN in existing quality assurance and supervisory models. It demonstrated originality in bringing these four frameworks together in a synthesized model for EQASUs in Thailand.

2.6. Identified Literature Gaps

Even though EQASUs play a crucial part in a national education quality assurance system (Mwinyipembe & Orodho, 2014), the requisite competencies that predict the effective functioning of an EQASU seldom attract the attention of many existing studies, especially in a non-Western context. Generally, studies deal with quality assurance at the institutional or systemic levels, rather than at the EQASU level. Very little research has been done on the study of supervisory units of quality assurance regarding the internal performance and training needs of EQASU staff, as well as the management structure of EQASU (Lucander & Christenson, 2020; Shawyun, 2012).

In addition, there are no agreed-upon global standards for the operation of EQASUs, resulting in piecemeal or incoherent models in many countries. In Thailand, previous EQA reforms have been criticized for their little effect on improving educational quality outcomes, partly due to a lack of validated performance indicators for EQASUs (Lee et al., 2023; Mala, 2018). The present study addresses the limited understanding by developing and validating a model specific to Thailand, which could also be adapted for similar educational systems.

In sum, including these components constructs an integrated framework for effective EQASUs functioning, corresponding to ONESQA's overarching goals of quality, creativity, participation, performance-centered, decentralization, flexibility and diversity, continuity, transparency and accountability, and approach and assurance. These foundational principles ensure that the EQA process remains relevant, credible, and impactful, ultimately contributing to the continuous improvement of Thailand's education system. This then leads to the authors' development of these research questions (RQ) and objectives (RO):

Cooperative networks are identified as a key to effective QA, promoting trust among QA actors and reducing redundant exercises among these actors (Kehm, 2010). Within the European Higher Education Area (EHEA), such cooperation has proven to foster greater convergence of standards and comparability of HE qualifications, and to build up legitimacy and trust amongst QA actors (Grifoll et al., 2015). Tightening networks and collaboration amongst EQASUs is thus likely to enhance the consistency and reliability of the QA system in Thailand.

Beyond the absence of EQASU-related studies, the current literature also attests to significant methodological limitations. Most of the studies reviewed in this research have used quantitative, cross-sectional survey approaches that favor inconsiderate numerical indicators over in-depth processual understanding. There is thus a lack of knowledge about how EQASUs operate in practice, how framing staff experience the supervisory process, or what contextual factors influence their effectiveness. In addition, previous studies rarely problematize divergences in definitions of supervision or the mismatch between indicators of quality and concrete performance. Bridging this gap calls for validated quantitative models and in-depth qualitative means that can sufficiently account for the lived experiences, institutional innerworkings, and cultural backgrounds defining various EQASUs. The present study represents an initial attempt to fill this gap by validating a competency-based framework and suggesting implications for further mixed-method research.

2.7. Research Questions (RQ)

RQ1. What are the critical components and indicators of quality for EQA supervisory units?

RQ2. How do these components and indicators, theorized through a structural-configurational lens (Mintzberg, 1980), align with the empirical practices observed in high-performing supervisory units?

RQ3. What are the key differences between high-performing and general EQA supervisory units regarding adherence to these components and indicators?

3. Methodology

3.1. Participants and Data Collection

This study employed a two-stage mixed-methods design to develop and validate the competency framework for External Quality Assessment Supervisory Units (EQASUs). In Stage 1, expert consultation was used to establish core components and indicators. In Stage 2, a large-scale survey and statistical modeling tested the framework's validity.

Participants in Stage 1 were nine experts purposively selected for their experience in external quality assessment, educational policy, and institutional supervision. In Stage 2, 434 external quality assessors participated, drawn from institutions across Thailand.

Data were analyzed with descriptive statistics, including frequency, percentage, mean, and standard deviation (SD). To rate the appropriateness of these components and indicators by the nine experts, a 5-point Likert scale questionnaire was used. Result interpretation used 4.51-5.00 (5) as most appropriate. All participants provided informed consent, and the study adhered to ethical research standards. The results in Table 1 show that the experts rated all four proposed components as 'Most Appropriate' (mean scores ranging from 4.56 to 4.67), providing strong evidence for their content and face validity.

Table 1. Expert Opinions on the Appropriateness of the Quality Components of EQASUs

Assessment Components	Experts (<i>n</i> =9)		Interpretation
	$\bar{x}$	SD	
Organizational Management (OM)	4.56	0.30	Highest
Personnel Performance Development (PPD)	4.67	0.21	Highest
Productivity and Results (PAR)	4.58	0.19	Highest
Cooperation Networks (CN)	4.60	0.42	Highest
Averages	4.70	0.33	Highest

3.2. Instrument Development

The initial draft instrument was developed based on a review of literature on quality assurance in education and supervisory practices, combined with the professional experience of the research team (Naz et al., 2022). A structured questionnaire was developed from Stage 1 findings, covering four hypothesized components: Organizational Management (OM), Personnel Performance Development (PPD), Productivity and Results (PAR), and Cooperative Networks (CN). Each component was represented by a set of observable indicators (see Appendix I, Table A1). These indicators were designed to capture measurable supervisory capacity aspects that align with international standards and Thai educational priorities.

3.3. Expert Validation (Stage 1: Delphi)

To establish content validity, a Delphi process was conducted with 18 experts (Lee et al., 2024). Using a structured questionnaire, participants rated each indicator's clarity, relevance, and sufficiency (Sottiyotin et al., 2023). Data were analyzed with descriptive statistics, including frequency, percentage, mean, and standard deviation (SD). To rate the appropriateness of these components and indicators, a 5-point Likert scale questionnaire was used. The tool was examined for content validity, and its IOC was between 0.80 and 1.00 (Praneetpolkrung & Supakicco, 2023). Result interpretation used the following criteria: 4.51-5.00: most appropriate, 3.51-4.50: highly appropriate, 2.51-3.50: moderately appropriate, 1.51-2.50: less appropriate, 1.00-1.50: least appropriate.

The interview guide was refined according to the experts' suggestions. Data were collected through individual in-depth Zoom interviews, which were recorded. The interview results were transcribed and synthesized before sending the results back to the interviewees to verify the accuracy of the data. The collected data were analyzed, compared, and synthesized to obtain a preliminary summary of the components and quality indicators of the EQASUs. The experts' consensus confirmed the appropriateness of the four-component model, and minor adjustments were made to indicator wording. Ratings of appropriateness are reported in the Appendix, Table A2.

3.4. Data Collection and Analysis (Stage 2)

In 2023, the revised instrument was administered to 434 external quality assessors (EQAs). The sample was selected using a proportional stratified random sampling method, with the 37 EQASUs throughout Thailand serving as the strata, to ensure representation across all supervisory units. Two validated instruments developed in the study were used to collect perceptions on EQASU component alignment and performance. Data collection was coordinated through ONESQA, which oversees the EQA units (ONESQA, 2013, 2023). The researchers distributed the online Google Forms questionnaires to the external assessors via the EQA units. The data collection process took four weeks in August 2023. Data were analyzed in several steps:

- **Confirmatory Factor Analysis (CFA):** CFA tested the validity of the four-component structure. Model fit was evaluated using commonly reported indices (χ^2/df , RMSEA, CFI, TLI, SRMR).
- **Conducting a try-out of the questionnaire** with 30 external assessors from the year 2023, who were not part of the sample group. The discrimination power ranged from 0.45 to 0.67, and the reliability was between 0.78 to 0.90.
- **Reliability Analysis:** Construct reliability and average variance extracted (AVE) were calculated for each component.
- **Group Comparisons:** A Mann–Whitney U test compared the alignment of high-performing EQASUs with the validated model against that of the general group, with significance set at $p < 0.01$.

All analyses were conducted using standard statistical packages. Fit indices and model performance results are presented in the Results section.

3.5. Goodness-of-Fit (GoF) Evaluation

The model's fit was assessed with widely recognized indices: Chi-Square/df ratio, Comparative Fit Index (CFI), Goodness-of-Fit Index (GFI), and Root Mean Square Error of Approximation (RMSEA). Conventional thresholds ($\text{CFI} \geq 0.90$, $\text{RMSEA} \leq 0.06$) were applied (Byrne, 2013; Hu & Bentler, 1999; Pimdee et al., 2024). to evaluate the adequacy of the CFA model. Convergent validity was further assessed using factor loadings, composite reliability, and average variance extracted (AVE).

3.6. Meta-Evaluation Using the CIPP Framework

To evaluate the quality of the proposed EQASU framework and resulting policy recommendations, a meta-evaluation was conducted with eight independent experts using Stufflebeam's CIPP model (Context, Input, Process, Product). The framework emphasizes utility, feasibility, propriety, and accuracy as dimensions of evaluation quality (Stufflebeam, 2015; Toosi et al., 2021; Wongrugsu et al., 2022; Yarbrough et al., 2010). Ratings were collected via questionnaire and analyzed for central tendency and dispersion.

3.7. Ethics and Software

The study adhered to ONESQA's ethical standards. Participation was voluntary, informed consent was obtained from all respondents, and no personal identifiers were collected. Data analyses were conducted using LISREL 9.1 and Python 3.8 packages for structural modeling and statistical validation.

4. Results

4.1. External Quality Assessor Characteristics

Table 2 presents the demographic and professional background of the 434 participating EQAs. The majority were female (71.89%). Regarding age, 33.41% were in the 41–50-year group (the largest category), followed by 28.57% aged 61 years and above, 26.03% aged 50–60, and 11.99% aged 35–40. Most assessors (71.65%) were engaged in basic education, with 28.35% working in early childhood education. Regarding experience, 39.17% had 6–10 years, 30.88% had more than 10 years, and 29.95% had 1–5 years of experience.

Table 2. General Characteristics of the Sample Group of EQAs

Characteristics of the sample group	Assessors	%
Gender		
Men	122	28.11
Women	312	71.89
Age		
35-40	52	11.99
41-50	145	33.41
50-60	113	26.03
61 and up	124	28.57
Assessment Level		
Early Childhood	123	28.35
Basic Education	311	71.65
Evaluation experience		
1-5 years	130	29.95
6-10 years	170	39.17
Ten years and above	134	30.88

4.2. CFA Results for EQASU Indicators

The model demonstrated strong fit indices: $\chi^2 = 135.66$, $df = 142$, $p = 0.63$, $RMSEA = 0.00$, $CFI = 1.00$, $GFI = 0.97$, indicating an excellent fit to the data (Figure 1). To establish convergent validity, we examined the factor loadings, Average Variance Extracted (AVE), and Composite Reliability (CR) (Pimdee et al., 2024). The AVE values for all first-order components ranged from 0.43 to 0.62, and Composite Reliability (CR) values ranged from 0.83 to 0.87 (Table 3). According to psychometric standards, CR values above 0.70 indicate good internal consistency, and AVE values above 0.50 are considered excellent. While the AVE for OM (0.43) and PAR (0.45) are slightly below the 0.50 threshold, this is acceptable given their high CR values (>0.80), as CR is a more reliable measure of internal consistency in such cases. The second-order construct (Component Result) showed excellent reliability and convergence, with an AVE of 0.73 and a CR of 0.91.

The path coefficients (β) between the second-order factor and the first-order components were all significant, with PPD emerging as the strongest predictor ($\beta = 0.94$), followed by OM ($\beta = 0.88$), PAR ($\beta = 0.94$), and CN ($\beta = 0.73$). The structure of this validated second-order model is presented in Figure 1.

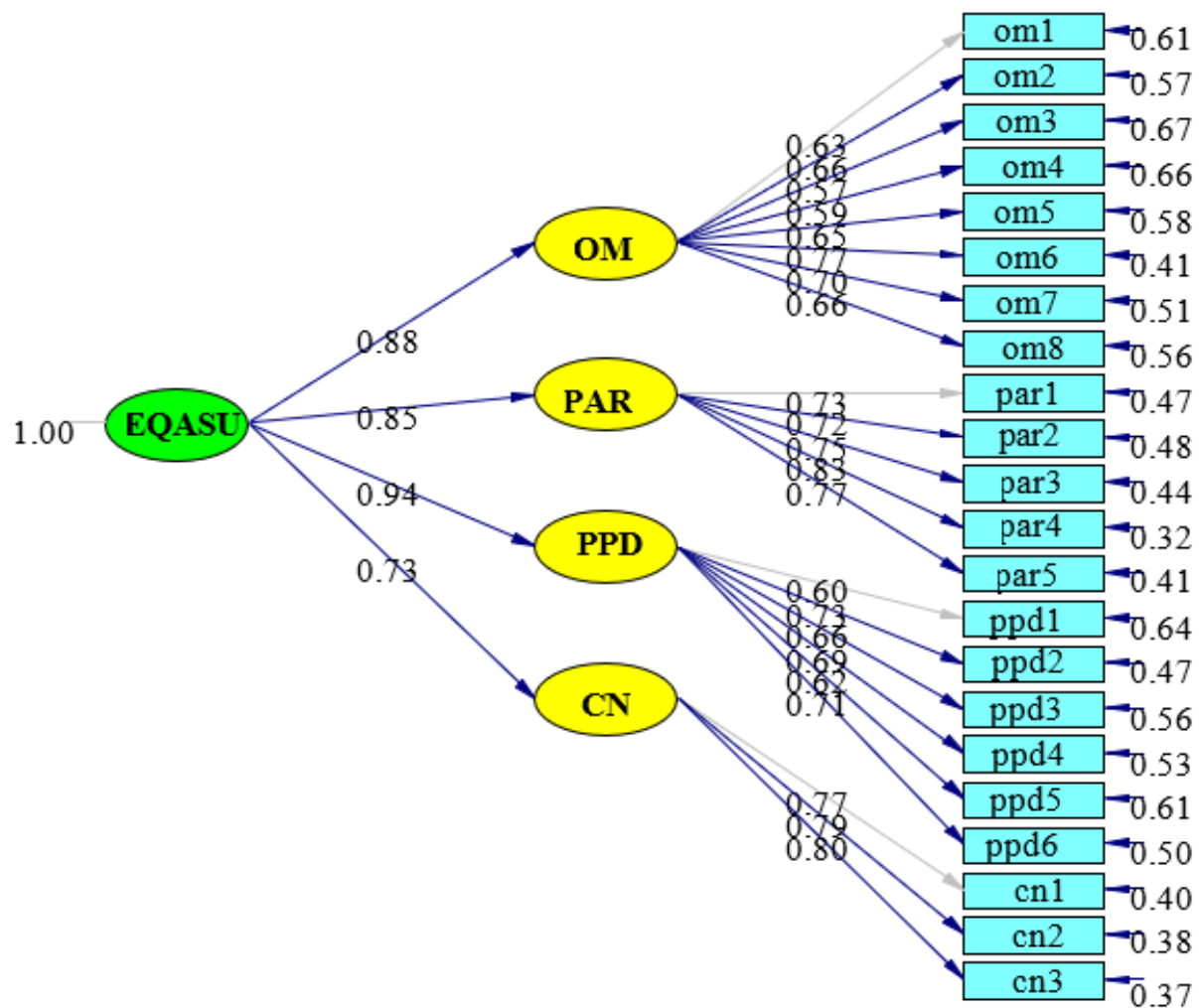


Figure 1. CFA Model of Standardized Component and Indicator Weights (β) for EQASU

Note. $\chi^2 = 135.66$, $df = 142$, $p = 0.63$, $RMSEA = 0.00$.

Source: The Authors' use of LISREL 9.1.

Table 3. First-Order Component AVEs and CRs

Components	AVE	Composite Reliability (CR)
Organizational Management (OM)	0.43	0.86
Personnel Performance Development (PPD)	0.58	0.87
Productivity and Results (PAR)	0.45	0.83
Cooperation Network (CN)	0.62	0.83
Component Results	0.73	0.91

4.3. Effectiveness of EQASU Components and Indicators

Table 4 summarizes the reliability and indicator weights for each component. PPD received the highest loading, highlighting the centrality of professional development for external assessors. This was closely followed by OM, CN, and PAR, confirming that high-performing EQASUs demonstrate alignment with all four validated domains. As shown in Table 4, all standardized factor loadings for the individual indicators were statistically significant ($p < .01$) and substantial, ranging from 0.57 to 0.83. This indicates that all indicators are strong measures of their respective latent constructs.

Table 4. Component Weights, Reliability of Observed Variables, and Indicator Coefficients for EQASU

Components and Indicators	(α)	β (SE)	(t)	(R ²)
Organizational management (OM)	0.81	0.88 (0.07)	11.93**	0.78
OM1: The organization has a systematic management structure that emphasizes participation.		0.63	<-->	0.39
OM2: The organization can manage its internal operations systematically and emphasize participation.		0.66	10.54**	0.43
OM3: The organization has a vision and strategies that promote work efficiency in alignment with the policies of ONESQA.		0.57	10.01**	0.33
OM4: There is an information system related to EQA that is convenient, comprehensive, and secure.		0.59	9.47**	0.34
OM5: There is a system for supervising and controlling the quality of internal quality assessment reports within the organization.		0.65	10.52**	0.42
OM6: Coordination between the supervisory unit and external quality assessors is conducted collegially.		0.76	11.20**	0.59
OM7: EQA planning aligns with the timelines set by ONESQA.		0.70	11.06**	0.49
OM8: The organization fosters a culture that emphasizes efficiency in EQA.		0.66	10.60**	0.44
Personnel performance development (PPD)	0.90	0.85 (0.06)	14.20**	0.72
PPD1: The organization strengthens the necessary competencies for EQA to meet the needs of the report reviewers and individual external assessors.		0.73**	<-->	0.53
PPD2: External assessors are guided and monitored to ensure their duties are performed according to the standards, criteria, and methods set by ONESQA with quality.		0.72**	13.79**	0.52
PPD3: There is a system for inspecting the EQAs conducted by individual external assessors.		0.75**	14.40**	0.56

Components and Indicators	(α)	β (SE)	(t)	(R ²)
PPD4: There are processes or methods to support or advise external assessors, enabling them to resolve various issues.		0.83**	15.71**	0.68
PPD5: Personnel within the supervisory unit are trained to have the knowledge and capability to review reports systematically.		0.77**	14.76**	0.59
Productivity and results (PAR)	0.78	0.94 (0.08)	11.68**	0.89
PAR1: The EQA reports are submitted within the timeline set by ONESQA.		0.60	<-->	0.36
PAR2: The EQA reports meet the criteria and principles established by ONESQA.		0.73	11.96**	0.53
PAR3: The EQA reports and accompanying documents are complete according to the components specified by ONESQA.		0.66	11.44**	0.44
PAR4: The EQA reports are accepted by educational institutions.		0.69	10.49**	0.47
PAR5: The EQA reports are valuable and relevant to the context of the educational institutions.		0.62**	9.82**	0.39
PAR6: The satisfaction of stakeholders within the supervisory unit is achieved.		0.71	10.72**	0.50
Cooperation networks (CN)	0.79	0.73 (0.06)	12.50**	0.54
CN1: There are effective processes or steps for communication and coordination with ONESQA.		0.77	<-->	0.60
CN2: Effective processes or steps for communication and coordination with other organizations exist.		0.79	15.36**	0.62
CN3: Support is provided for information and data related to EQA that aligns with ONESQA's policies.		0.80	15.56**	0.63
Chi-Square=135.66, df=142, p=0.63, $\chi^2/df=0.95$, RMSE=0.00, RMR=0.02, CFI=1.00, GFI=0.97, AGFI=0.95				

** $p < .01$, β = standardized component weights, <--> means forced parameters; therefore, SE and t values are not reported.

4.4. Meta-Evaluation of Policy Recommendations

The policy recommendations derived from the validated framework were subjected to a meta-evaluation by eight experts. As shown in Table 5, the overall average score was exceptionally high ($\bar{x}$ = 4.84, SD = 0.10). Utility scored the highest ($\bar{x}$ = 4.88), followed by feasibility ($\bar{x}$ = 4.85), accuracy ($\bar{x}$ = 4.83), and propriety ($\bar{x}$ = 4.80). These findings confirm that the recommendations were practical, ethical, and feasible. They also align with prior applications of Stufflebeam's CIPP model in

educational program evaluation, which emphasize the importance of utility, feasibility, accuracy, and propriety (Stufflebeam, 2015; Toosi et al., 2021; Wongrugsu et al., 2022; Yarbrough et al., 2010).

Table 5. Expert Evaluation Results of EQASU Policy Recommendations ($n = 8$)

Items	$\bar{x}$	SD	Interpretation
Utility	4.88	0.13	Most appropriate
Feasibility	4.85	0.24	Most appropriate
Propriety	4.80	0.31	Most appropriate
Accuracy	4.83	0.25	Most appropriate
Total average	4.84	0.10	Most appropriate

Note. 4.51-5.00 = *Most appropriate*.

Overall, the results confirm a valid and strong four-component framework for EQASUs. PPD is determined to be the most important dimension, with Organizational Management, Cooperative Networks, and Productivity resulting in significance. The CFA model also demonstrated a high goodness-of-fit statistic, indicating that the indicators appropriately captured the structure of effective supervisory units. The meta-evaluation by experts demonstrated the utility, feasibility, accuracy, and propriety of the policy recommendations, indicating that the validated framework was defensible in practice and theory. These results together provide a firm grounding for improving the quality of external educational assessments in Thailand or similar settings. These findings are directly related to the purpose of the research and show that the validity of the proposed EQASU framework is confirmed, and its components and indicators are empirically reliable and also applicable for professional and practical guidance of supervisors in the context of educational quality assessment.

5. Discussion

5.1. Organizational Management (OM)

Organizational management in the EQA process has a good reliability value ($\alpha = 0.81$) and a high effect ($R^2 = 0.78$) on the overall effectiveness of the EQA process. High t-values of all indicators (9.47–11.93) show their high contribution to the effectiveness of OM. Similarly, OM1 (*The organization has a systematic management structure emphasizing participation*), which has a moderate effect ($R^2 = 0.39$) and is essential to participatory management, ensuring that all stakeholders are included in the process of quality assurance (Lucander & Christenson, 2020). Alternatively, OM3 illustrates the importance of organizational vision and strategy alignment with the policies of ONESQA, which demonstrates the importance of strategic alignment in achieving educational quality ($R^2 = 0.33$). The impact of OM6 on coordination ($R^2 = 0.59$) introduces the possibility that the collegiality of the supervisory function vis-à-vis external assessors, in coordination with others in quality assurance, plays an essential role in the maintenance of the validity and reliability of assessments (Chaiya & Ahmad, 2021; Valentine et al., 2021; Weldearegay, 2020).

5.2. Personnel Performance Development (PPD)

The analysis reveals that Personnel Performance Development (PPD) is a fundamentally critical component for EQASU effectiveness, demonstrated by its high composite reliability ($\alpha = 0.90$) and its substantial explanatory power within the model ($R^2 = 0.72$). This indicates that the competencies and performance of personnel are not only highly consistent measures but also account for a significant majority of the variance in supervisory unit performance.

The findings particularly underscore the paramount importance of providing continuous support to assessors. The indicator PPD4 ('processes to support or provide advice to external assessors') emerged as the strongest independent predictor in its category ($R^2 = 0.68$). This empirically validates that moving beyond one-off training to establish ongoing mentoring, advice, and calibration mechanisms is essential for improving assessment quality and consistency (Dart et al., 2021; Matthews & Dobbins, 2021; Sato et al., 2008).

This can be theorized through Mintzberg's (1980) framework: a practical EQASU functions as a Professional Bureaucracy, which relies on standardized skills—achieved through robust PPD—as its core coordinating mechanism. The assessors' expert judgment is the organization's primary asset. Therefore, deficits in PPD directly undermine the entire system's credibility. For example, criticisms of inconsistent scoring between assessment teams in previous Thai EQA rounds can be directly traced to a lack of standardized, continuous personnel development and calibration (ONESQA, 2023). While Organizational Management, Productivity, and Networks are vital enabling factors, Personnel Performance Development is the engine that drives the quality and legitimacy of the core assessment work.

5.3. Productivity and Results (PAR)

PAR posed a high coefficient alpha ($\alpha = 0.78$) and significantly contributed to each EQA's effectiveness ($R^2 = 0.89$). This suggests that timely and accurate reports are fundamentally required in the quality assessment (Shawyun, 2012). PAR2 (*The EQA reports meet the criteria and principles established by ONESQA*) has a substantial impact ($R^2 = 0.53$), suggesting that adherence to ONESQA's (2013, 2023) criteria is a critical determinant of assessment quality. The findings also highlight the importance of stakeholder satisfaction (PAR6), which has a significant impact ($R^2 = 0.50$) on the perceived success of the assessment process.

5.4. Cooperative Networks (CN)

Finally, CN illustrates good reliability ($\alpha = 0.79$) and overall effect ($R^2 = 0.54$). The high significance of CN3 (*Support is provided for information and data related to EQA that aligns with ONESQA's policies*) ($R^2 = 0.63$) indicates that data sharing and effective communications entail necessary practices for maintaining conformity with national standards (Sempere et al., 2009). Likewise, other studies have remarked that effective communications create a collaborative environment and ensure the successful implementation of quality assessments (Awasthy et al., 2020; ONESQA, 2013).

The high reliability of the components and their significant impacts can also be determined from the standardized beta (β), which functions like the correlation coefficient, ranging from 0 - 1 or 0 - -1, depending on the direction of the relationship, and the closer a value is to 1 or -1, the stronger the relationship. Hence, it can be deduced that the primary construct PPD was most important ($\beta = 0.94$). This was then followed from a distance by OM ($\beta = 0.81$), CN ($\beta = 0.79$), and PAR ($\beta = 0.78$).

These findings align with Sato et al. (2008)'s study, which also stressed strategic alignment, ongoing support, and communication as crucial success factors conducive to 'achieving high-quality educational assessment.' Assessor confidence or self-efficacy has also been identified as an important factor in encouraging uptake and engagement with learning technologies (Zheng et al., 2020). These findings indicate that organizations involved in EQAs should focus on developing management structures, providing continuous assessor support (Matthews & Dobbins, 2021), and building strong communication networks. Observing what has been happening in other countries, Lee et al. (2023) have suggested and forewarned that large-scale educational assessment testing in some countries is regarded as a waste of school time while missing potential new teaching and learning methods. This has been an important issue in Thailand's case and was a primary reason for delaying the fourth testing round by two years (Sriarunrasmee & Seekheio, 2024). Further research could investigate how

these components interact and how they can be further enhanced to improve the overall quality assurance process.

5.5. Results of the Verification of Validity of Components and Quality Indicators for EQASUs

The results presented in Table 6 highlight a significant distinction between the performance of External Quality Assessment Supervisory Units (EQASUs) in the High-Performance Group (HPG) and the General Group (GG). The analysis shows that the HPG demonstrates the highest level of adherence to the key quality components, while the GG only achieves a "high" level of adherence. This difference underscores the superior implementation of quality standards in the HPG, suggesting that the components are more effectively integrated and practiced within this group.

Table 6. Results of Verifying the Validity of Components and Quality Indicators for EQASUs

Components	Groups						Mann-Whitney U Test	Sig
	High-Performance Group			General Group				
	Mean	SD	Level	Mean	SD	Level		
Organizational Management (OM)	4.72	0.11	Highest	3.90	0.25	High	2.611**	.00
Personnel Performance Development (PPD)	4.82	0.10	Highest	3.89	0.46	High	2.611**	.00
Productivity and Results (PAR)	4.82	0.12	Highest	4.00	0.31	High	2.619**	.00
Cooperation Network (CN)	4.50	0.18	Highest	3.68	0.41	High	2.611**	.00
Component Results	4.72	0.04	Highest	3.88	0.32	High	2.611**	.00

Note. ** $p < .01$

These findings were further confirmed with statistical support from the Mann-Whitney U Test, a non-parametric method of comparing groups. This was ideal in this case as it does not make the assumption of normality about the distributions of data, and found that HPG's proportions adhering to OM, PPD, PAR, and CN were significantly higher than GG's proportions adhering to the same four components at the 0.01 level of statistical significance. This means there is a less than 1% probability that these differences in proportions are derived from random variation rather than systematically better adherence of HPG compared to GG for each component.

These findings are in the justification of the suggested quality components for EQASUs. The noticeable statistically significant differences between HPG and GG in all components show that the framework for analyzing these units is credible and valid, as the low U values and small p -values of the Mann-Whitney U Test suggest that the HPG performs consistently 'better' than the GG, by indicating that the quality indicators can differentiate high-performing from general units (Riina et al., 2023). Based on prior research (Riina et al., 2023), the Mann-Whitney U Test is similar to that of a t-test, however as opposed to a t-test of which it assumes a normal distribution, the Mann-Whitney U Test makes no assumption of this kind, making it more suitable to this analysis as perhaps the distribution of this data may follow a different distribution.

These findings are significant because they validate the quality components and indicators for EQASUs, thereby verifying that the framework is theoretically valid and practically applicable in reality. Consequently, the components such as organizational management and personnel development are not merely nominal elements, but they are operational features that can notably improve the quality of educational assessments when they are put into practice. The higher performance of the HPG also indicates that the higher degree of compliance with the validated components leads to the higher achievement of EQASUs, which is a good roadmap for other quality supervisory units to improve their performance.

Finally, the results corroborate the validity, strength, and stability of the EQASU quality framework. The statistically significant differences identified between the HPG and GG highlight the sufficiency of the proposed components in accurately measuring quality. This finding underscores the framework's efficacy in guiding EQASUs toward improved performance and outcomes.

5.6. Synthesis and Implications

This study developed and validated a holistic framework for External Quality Assessment Supervisory Units (EQASUs), identifying four core components, including Organizational Management (OM), Personnel Performance Development (PPD), Productivity and Results (PAR), and Cooperative Networks (CN), as essential to their effectiveness. A CFA was performed to empirically validate the strength of the second-order CFA structure, with PPD ($\beta = 0.94$) emerging as the most influential component, followed by OM ($\beta = 0.81$), CN ($\beta = 0.79$), and PAR ($\beta = 0.78$). Notably, the known-groups analysis confirmed the framework's practical validity, demonstrating that high-performing EQASUs adhere to these components significantly more than their counterparts.

When interpreted alongside prior research, this framework offers a nuanced contribution. While international models like EFQM and Baldrige provide valuable macro-level guidance on quality management (Brennan & Shah, 2000; Harvey & Newton, 2007), our findings operationalize these principles for the specific, meso-level context of the *supervisory unit*. For instance, the paramount importance of PPD empirically validates calls for "investing in assessors" (Kohoutek & Westerheijden, 2014). It provides a measurable blueprint for how to do it, moving beyond abstract principle to defined practice (e.g., through mechanisms like PPD4's ongoing support). Similarly, the strong weighting of OM and CN underscores that effective quality assurance is not merely a technical process but a socio-structural one, dependent on participatory management and collaborative networks to achieve legitimacy and avoid being perceived as a mere "compliance exercise" (Brady & Bates, 2016; Krooi et al., 2024; Stensaker, 2008).

The unique contribution of this study is threefold. First, it addresses a critical gap in the literature by providing the first empirically validated competency framework designed explicitly for EQASUs, moving beyond institutional-level QA to focus on the bodies that ensure its credibility. Second, it resolves ambiguities in the Thai context and similar systems by offering clear, evidence-based indicators to define and standardize effective supervisory practice, directly addressing historical criticisms of inconsistency and unclear roles (Mala, 2018; Saengpassa, 2014). Finally, it theorizes the core function of an EQASU through Mintzberg's lens, arguing convincingly that it must be structured as a professional bureaucracy where standardizing assessor skills (PPD) is the primary mechanism for ensuring quality, not just standardizing processes (PAR).

For policymakers and practitioners, the implications are clear. Investment in EQASUs must prioritize continuous, structured personnel development above all else. Building a mentoring, calibration, and support system is not an optional add-on but the fundamental engine of assessment quality and credibility. This research provides ONESQA and other national agencies with a validated roadmap to strengthen their supervisory structures, enhance the legitimacy of their QA outcomes, and ultimately, fulfill their mission of driving meaningful educational improvement.

6. Conclusion

The study successfully developed and validated a comprehensive model for assessing the quality of EQA units. The identified components—Organizational Management, Personnel Development, Productivity and Results, and Cooperative Networks—proved theoretically sound and empirically significant. The significant differences observed between high-performance and general groups validate the importance of these components in achieving high-quality outcomes. The study provides an empirically validated framework where the critical role of continuous personnel development is clearly demonstrated as the foremost driver of supervisory unit effectiveness. This research contributes a validated model that can serve as a benchmark for future assessments, ensuring the consistent and effective evaluation of quality assessment supervisory units.

7. Recommendations for EQASUs

ONESQA should develop strategies that emphasize the importance of the identified components and quality indicators in selecting and evaluating EQASUs. This could include:

- Develop a selection process that prioritizes these components and assigns appropriate weight to each indicator during evaluation.
- Encourage systems that allocate tasks based on the readiness and potential of the EQA units.
- Foster the implementation of internal quality assurance systems within EQASUs to maintain high standards consistently (Ra-ngubtook & Bhongsatiern, 2022).
- Develop a certification process that evaluates and certifies EQASUs based on their quality and capacity, with certification periods ranging from 2 to 5 years.

ONESQA should also implement measures to ensure the practical application of these strategies:

- Appoint working groups, organize meetings, and establish operational guidelines to ensure all stakeholders—such as ONESQA staff, external assessors, and quality assessment unit managers—understand the components and quality indicators.
- Produce promotional materials to engage relevant stakeholders in the planned activities.
- Implement the planned activities and survey participants to assess their understanding of the components and quality indicators, and their satisfaction with the process.
- Create various types of documentation related to the components and quality indicators to be used by stakeholders in fulfilling their roles and responsibilities.
- Evaluating the Utility, Feasibility, propriety, and accuracy of the implemented strategies should result in scores at or above 3.50, indicating successful application.

Declarations

Credit Authorship Contribution Statement. Aukkapong Sukkamart (AS), Jaruwan Ployduangrat (JP1), Jakkapong Prongprommarat (JP2), and Paitoon Pimdee (PP) Conceptualization, AS and PP; Software, PP, JP2; Validation, JP1, AS; Formal analysis, AS, PP; Investigation, JP1 and AS; Resources, AS and JP1; Writing—original draft preparation, AS and JP2; Writing—review and editing, JP1 and AS; Supervision, AS, PP; All authors have read and agreed to the published version of the manuscript. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest. The authors declare no conflict of interest.

Funding. This research project was funded by the Office for National Education Standards and Quality Assessment (Public Organization) (ONESQA) (Grant No: NHCT.2566/40).

Ethical Approval. This research did not engage with vulnerable groups, collect identifiable personal information, or implement any intervention. Consequently, in accordance with Thailand's *Guidelines for Conducting Human Subjects Research in Behavioral Science, Social Sciences, and Humanities* (Phuangsuwan et al., 2024), the study was exempt from the requirement for formal ethics review. An exemption waiver was also issued by the authors' university ethics committee in compliance with internationally recognized ethical standards, including the *Declaration of Helsinki*, the *Belmont Report*, the *CIOMS Guidelines*, the *International Conference on Harmonization – Good Clinical Practice (ICH-GCP)*, and *45 CFR 46.101(b)* of the U.S. Federal Regulations. As the study was classified as exempt, no continuing review was required. Finally, Informed consent was obtained from all participants involved in the study.

Data Availability Statement. The data supporting the findings of this study are available from the corresponding author upon reasonable request. All datasets were collected and stored in compliance with institutional research ethics and privacy guidelines.

Acknowledgements. The researchers would like to thank the experts who contributed invaluable suggestions for developing the components and quality indicators for the EQA units. Special thanks are also extended to the specialists who carefully reviewed the research instruments and to the external assessors who participated as respondents in this study. The authors also thank Ajarn Charlie for his invaluable contribution in editing and proofreading this paper. In preparing this manuscript, the authors utilized Scholar GPT and Google Translate to assist in translating the work from Thai to English. Grammarly AI Premium was also used to review grammar and structure before submission. In addition to using AI tools for translation and language correction, all the content underwent a thorough review and editing by a native English speaker working in collaboration with the authors. Therefore, the authors assert full responsibility for the final manuscript's accuracy, originality, and quality.

References

- Alam, L., Alam, M., Shafi, M. N. U. H., Khan, S., & Khan, Z. M. (2022). Meaningful in-training and end-of-training assessment: The need for implementing a continuous workplace-based formative assessment system in our training programs. *Pakistan Journal of Medical Sciences*, 38(5), 1132. <https://doi.org/10.12669%2Fpjms.38.5.5921>
- Asampong, E., Kamau, E. M., Teg-Nefaah Tabong, P., Glozah, F., Nwameme, A., Opoku-Mensah, K., ... & Dako-Gyeke, P. (2023). Capacity building through comprehensive implementation, research, training, and mentorship: an approach for translating knowledge into practice. *Globalization and Health*, 19(1), 35. <https://doi.org/10.1186/s12992-023-00935-8>
- Ayeni, A. J. (2012). Assessment of Principals' Supervisory Roles for Quality Assurance in Secondary Schools in Ondo State, Nigeria. *World Journal of Education*, 2(1), 62-69. <https://eric.ed.gov/?id=EJ1158929>
- Awasthy, R., Flint, S., Sankarnarayana, R., & Jones, R. L. (2020). A framework to improve university-industry collaboration. *Journal of Industry-University Collaboration*, 2(1), 49-62. <https://doi.org/10.1108/JIUC-09-2019-0016>
- Behraves, E., Abubakar, A. M., & Tanova, C. (2021). Participation in decision-making and work outcomes: evidence from a developing economy. *Employee Relations: The International Journal*, 43(3), 704-723. <https://doi.org/10.1108/ER-08-2018-0228>
- Brennan, J., & Shah, T. (2000). Quality assessment and institutional change: Experiences from 14 countries. *Higher education*, 40(3), 331-349. <https://doi.org/10.1023/A:1004159425182>

- Brady, N., & Bates, A. (2016). The standards paradox: How quality assurance regimes can subvert teaching and learning in higher education. *European Educational Research Journal*, 15(2), 155-174. <https://doi.org/10.1177/1474904115617484>
- Burgess, M., & Phillips, H. (2021). Authentic assessments and the external examiner. In *The Role of External Examining in Higher Education: Challenges and Best Practices* (Vol. 38, pp. 43-59). Emerald Publishing Limited. <https://doi.org/10.1108/S2055-364120210000038004>
- Byrne, B. M. (2013). Structural equation modeling with Mplus: Basic concepts, applications, and programming. Routledge. <https://doi.org/10.4324/9780203807644>
- Castanelli, D. J., Weller, J. M., Molloy, E., & Bearman, M. (2022). How trainees come to trust supervisors in workplace-based assessment: a grounded theory study. *Academic Medicine*, 97(5), 704-710. <https://doi.org/10.1097/ACM.0000000000004501>
- Chaiya, C., & Ahmad, M. M. (2021). Success or failure of the Thai higher education development—Critical factors in the policy process of quality assurance. *Sustainability*, 13(17), 9486. <https://doi.org/10.3390/su13179486>
- Chikazinga, W. W. (2025). University Actors Responding to the National Quality Assurance Regime for Higher Education in Malawi: A Case of Instrumentality, Pragmatism and Symbolic Compliance. *Higher Education Quarterly*, 79(1), e12577. <https://doi.org/10.1111/hequ.12577>
- Child, S., & Ellis, P. (2021). *The what, why, and how of assessment: A guide for teachers and school leaders*. Sage Publications. <https://www.torrossa.com/en/resources/an/5018839>
- Dart, J., Twohig, C., Anderson, A., Bryce, A., Collins, J., Gibson, S., ... & Palermo, C. (2021). The value of programmatic assessment in supporting educators and students to succeed: A qualitative evaluation. *Journal of the Academy of Nutrition and Dietetics*, 121(9), 1732-1740. <https://doi.org/10.1016/j.jand.2021.01.013>
- Dill, D. D. (2020). Enhancing academic quality and collegial control: Insights from U.S. policy on the ethical conduct of human subjects' research. *Higher Education Policy*, 33(1), 45-64. <https://doi.org/10.1057/s41307-018-0093-9>
- Duarte, N., & Vardasca, R. (2023). Literature review of accreditation systems in higher education. *Education Sciences*, 13(6), 582. <https://doi.org/10.3390/educsci13060582>
- El-Khawas, E. (2014). Inside quality reform: Early results on using outcomes for improvement. *Quality in Higher Education*, 20(2), 183-194. <https://doi.org/10.1080/13538322.2014.926083>
- Galeazzo, A., Furlan, A., & Vinelli, A. (2021). The role of employees' participation and managers' authority on continuous improvement and performance. *International Journal of Operations & Production Management*, 41(13), 34-64. <https://doi.org/10.1108/IJOPM-07-2020-0482>
- Geisinger, K. F. (1994). Cross-cultural normative assessment: translation and adaptation issues influencing the normative interpretation of assessment instruments. *Psychological Assessment*, 6(4), 304. <https://doi.org/10.1037//1040-3590.6.4.304>
- Grifoll, J., Hopbach, A., McClaran, A., Valeikiene, A., Sanchez, T., & Ranne, P. (2015). Quality Procedures in the European higher education area and beyond—internationalisation of Quality assurance agencies. 4th ENQA survey. *Brussels: ENQA*. https://www.enqa.eu/wp-content/uploads/enqa_oc_22.pdf
- Harris, P., Bhanji, F., Topps, M., Ross, S., Lieberman, S., Frank, J. R., ... & ICBME Collaborators. (2017). Evolving concepts of assessment in a competency-based world. *Medical Teacher*, 39(6), 603-608. <https://doi.org/10.1080/0142159X.2017.1315071>

- Harvey, L., & Newton, J. (2007). *Transforming quality evaluation: moving on. In Quality assurance in higher education: Trends in regulation, translation and transformation* (pp. 225-245). Springer Netherlands. <https://doi.org/10.1007/978-1-4020-6012-0>
- Henderson, C. (2025). *Teacher wellbeing and the shaping of teacher shortages in crisis-affected contexts*. Paper commissioned for the 2024 Global Report on Teachers. <https://tinyurl.com/4uh24fej>
- Hu, L. T., & Bentler, P. M. (1999). Cut-off criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1 – 55. <http://doi.org/10.1080/10705519909540118>
- Jöreskog, K. G., Olsson, U. H., & Wallentin, F. Y. (2016). *Multivariate analysis with LISREL*. Springer. <https://doi.org/10.1007/978-3-319-33153-9>
- Kaiser, F., Melo, A. I., & Hou, A. Y. C. (2022). Are quality assurance and rankings useful tools to measure 'new' policy issues in higher education? The practices in Europe and Asia. *European Journal of Higher Education*, 12(sup1), 391-415. <https://doi.org/10.1080/21568235.2022.2094816>
- Kehm, B. M. (2013). To Be or Not to Be? The Impacts of the Excellence Initiative on the German System of Higher Education. In Shin, J., Kehm, B. (Eds.), *Institutionalization of World-Class University in Global Competition. The Changing Academy – The Changing Academic Profession in International Comparative Perspective*, vol 6. Springer. https://doi.org/10.1007/978-94-007-4975-7_6
- Kohoutek, J., & Westerheijden, D. (2014). 13. Opening up the black box. In Eggins, H. (Ed.), *Drivers and barriers to achieving quality in higher education*, 167. https://doi.org/10.1007/978-94-6209-494-9_13
- Krooi, M., Whittingham, J., & Beusaert, S. (2024). Introducing the 3P conceptual model of internal quality assurance in higher education: A systematic literature review. *Studies in Educational Evaluation*, 82, 101360. <https://doi.org/10.1016/j.stueduc.2024.101360>
- Lee, J., Lee, J. S., & Ellis, N. (2023). Public Opinion about National Large-Scale Student Assessment: A case of NAPLAN. *Educational Assessment*, 28(3), 137-155. <https://doi.org/10.1080/10627197.2023.2215977>
- Lee, U., Jung, H., Jeon, Y., Sohn, Y., Hwang, W., Moon, J., & Kim, H. (2024). Few-shot is enough: exploring ChatGPT prompt engineering method for automatic question generation in English education. *Education and Information Technologies*, 29(9), 11483-11515. <https://doi.org/10.1007/s10639-023-12249-8>
- Leiber, T., Stensaker, B., & Harvey, L. (2015). Impact evaluation of quality assurance in higher education: methodology and causal designs. *Quality in Higher Education*, 21(3), 288-311. <https://doi.org/10.1080/13538322.2015.1111007>
- Loukkola, T., & Zhang, T. (2010). *Examining quality culture: Part 1-Quality assurance processes in higher education institutions*. Brussels: European University Association. <https://aic.lv/bolona/2010/Reports/Examining%20Quality%20Culture%20Part%201.pdf>
- Lucander, H., & Christenson, C. (2020). Engagement for quality development in higher education: A process for quality assurance of assessment. *Quality in Higher Education*, 26(2), 135-155. <https://doi.org/10.1080/13538322.2020.1761008>

- Ma, M. (2023). *Exploring employee participation in corporate volunteering in China: Affecting factors and outcomes* (Doctoral dissertation, Hong Kong Baptist University). China. <https://tinyurl.com/bdef56zv>
- Mala, D. (2018, August 11). After delays, Onesqa to begin evaluations. *Bangkok Post*. <https://tinyurl.com/y3jnxvvt>
- Mutluer, C. (2023). Development of the scientific literacy scale: A study of validity and reliability. *International Journal of Educational Researchers*, 14(3). <https://doi.org/10.1016/10.29329/ijer.2023.595.4>
- Mwinyipembe, M. M., & Orodho, A. J. (2014). Effectiveness of quality assurance and standards officers school supervisory roles in enhancing students' academic performance in national examinations in Nakuru District, Kenya. *Journal of Education and Practice*, 5(16), 69-80. <https://tinyurl.com/2vy5cktb>
- Nadeem, M. (2023). Leadership's Role in Effective Quality Assurance Implementation in Higher Education Institutions. *Education Quarterly Reviews*, 6(3). <https://doi.org/10.31014/aior.1993.06.03.778>
- Narayanamma, P. L., & Lalitha, K. (2016). Balanced scorecard-The learning & growth perspective. *Aweshkar Research Journal*, 21(2), 59.
- Naz, N., Gulab, F., & Aslam, M. (2022). Development of qualitative semi-structured interview guide for case study research. *Competitive Social Science Research Journal*, 3(2), 42-52. <https://cssrjournal.com/index.php/cssrjournal/article/view/170>
- Newton, J. (2010). A tale of two 'qualitys': reflections on the quality revolution in higher education. *Quality in Higher Education*, 16(1), 51-53. <https://doi.org/10.1080/13538321003679499>
- Nguyen, Q. L. H. T. T., CHU, N. N. M., & Van Hong, T. R. A. N. (2020). Application of total quality management in developing quality assessment model: The case of Vietnamese higher education. *The Journal of Asian Finance, Economics and Business*, 7(11), 1049-1057. <https://doi.org/10.13106/jafeb.2020.vol7.no11.1049>
- ONESQA. (2013, October). *Manual for Educational Institutions: The third round of external quality assessment (2011-2015)*. <https://tinyurl.com/39ndn6xs>
- ONESQA. (2017). *About ONESQA*. <http://www.onesqa.or.th/en/profile/973/>
- ONESQA. (2023, June 23). *Guidelines for the Fourth Round EQA for Higher Education*. <https://www.onesqa.or.th/en/profile/1291/>
- Pimdee, P., Sukkamart, A., Nantha, C., Kantathanawat, T., & Leekitchwatana, P. (2024). Enhancing Thai student-teacher problem-solving skills and academic achievement through a blended problem-based learning approach in online flipped classrooms. *Heliyon*, 10(7). <https://doi.org/10.1016/j.heliyon.2024.e29172>
- Pitiyanuwat, S. (2011). Quality assurance in South-East Asian higher education. In Brock, C., & Pe Symaco, L. (Eds.). *Education in South-East Asia*. <https://tinyurl.com/27n2tzf9>
- Pitiyanuwat, S., Phanphruk, S., & Pitiyanuwat, T. (2018). Educational testing, assessment, and quality assurance. In Fry, G. (Ed.). *Education in Thailand. Education in the Asia-Pacific Region: Issues, Concerns and Prospects*, vol 42. Springer. https://doi.org/10.1007/978-981-10-7857-6_24
- Pizzolato, D., & Dierickx, K. (2023). Research integrity supervision practices and institutional support: A qualitative study. *Journal of Academic Ethics*, 21(3), 427-448. <https://doi.org/10.1007/s10805-022-09468-y>

- Praneetpolkrung, K., & Supakicco, P. S. (2023). Empowering Teachers' Learning to Enhance Students' Change Leadership Skills. *World Journal of Education*, 13(3), 1-12. <https://doi.org/10.5430/wje.v13n3p1>
- Preston, R., Gratani, M., Owens, K., Roche, P., Zimanyi, M., & Malau-Aduli, B. (2020). Exploring the impact of assessment on medical students' learning. *Assessment & Evaluation in Higher Education*, 45(1), 109-124. <https://doi.org/10.1080/02602938.2019.1614145>
- Phuangsuwan, P., Limna, P., & Siripipatthanakul, S. (2024). Ethics in the Social Sciences Research. *Advance Knowledge for Executives*, 3(4), No. 49. <https://tinyurl.com/4a8b2e34>
- Ra-ngubtook, W., & Bhongsatiern, J. (2022). Basic education in Thailand. In *International Handbook on Education in South East Asia* (pp. 1-25). Springer Nature. https://doi.org/10.1007/978-981-16-8136-3_21-1
- Riina, M. D., Stambaugh, C., Stambaugh, N., & Huber, K. E. (2023). Continuous variable analyses: t-test, Mann–Whitney, and Wilcoxon rank. In *Translational Radiation Oncology* (pp. 153-163). Academic Press. <https://doi.org/10.1016/B978-0-323-88423-5.00070-4>
- Roengsumran, A. (2015). *A Study of External School Evaluators' Competencies for Grade Levels 1-12 Quality Assessment in Thailand* (Doctoral dissertation, University of Minnesota). Minnesota. <https://tinyurl.com/4zp986kt>
- Saengpassa, C. (2014, May 19). ONESQA director rejects the criticism. *The Nation*. <https://www.nationthailand.com/in-focus/30233944>
- Schwarz, S., & Westerheijden, D. F. (Eds.). (2007). *Accreditation and evaluation in the European higher education area (Vol. 5)*. Springer Science & Business Media.
- Sempere, A. G., Mohn, C., & Pieroni, R. (2009). Improving educational effectiveness and promoting internal and external information—sharing through student learning outcomes assessment. *Toward useful program evaluation in college foreign language education*, 57-81. <https://tinyurl.com/3fjwv4fe>
- Shawyun, T. (2012). Strategic performance management system: An integrated framework. *International Edition*. Center for Excellence, Assumption University of Thailand. <https://tinyurl.com/3684fzxv>
- Sottiyotin, T., Uitrakul, S., Sakdiset, P., Sukkarn, B., Sangfai, T., Chuaboon, L., & Palee, P. (2023). Effective formative assessment for pharmacy students in Thailand: Lesson learns from a school of pharmacy in Thailand. *BMC Medical Education*, 23(1), 300. <https://doi.org/10.1186/s12909-023-04232-1>
- Sriarunrasmee, J., & Seekheio, D. (2024). The development of a support system for the follow-up on basic education quality assessment and accreditation. *Educational Administration: Theory and Practice*, 30(7), 532-539. <https://doi.org/10.53555/kuey.v30i7.6717>
- Sriputtarin, S. (2008). Developing a model for certifying external assessors in educational quality assessment of basic education. *Education Journal of Thailand*, 2(1), 73-83. <https://ojs.lib.buu.ac.th/index.php/Education3/article/view/1630>
- Stensaker, B. R. (2008). Outcomes of quality assurance: A discussion of knowledge, methodology and validity. *Quality in Higher Education*, 14(1), 3-13. <https://doi.org/10.1080/13538320802011532>
- Stensaker, B., & Leiber, T. (2015). Assessing the organisational impact of external quality assurance: Hypothesising key dimensions and mechanisms. *Quality in Higher Education*, 21(3), 328-342. <https://doi.org/10.1080/13538322.2015.1111009>

- Stufflebeam, D. L. (2015). CIPP evaluation model checklist: A tool for applying the CIPP model to assess projects and programs. *Western Michigan University Evaluation Center*. <http://tinyurl.com/mucdxtm7>
- Tangwarasittichai, K., Leekitchwatana, P., & Pimdee, P. (2021). Thai early childhood education external assessor competency indicator development. *PalArch's Journal of Archaeology of Egypt/Egyptology*, 18(3), 120-132. <https://tinyurl.com/un2f6883>
- Toosi, M., Modarres, M., Amini, M., & Geranmayeh, M. (2021). Context, Input, Process, and Product Evaluation Model in medical education: A systematic review. *Journal of Education and Health Promotion*, 10(1), 199. https://doi.org/10.4103/jehp.jehp_1115_20
- United Nations Educational, Scientific, & Cultural Organization. (2016). *Reviews of National Policies for Education in Thailand An OECD-UNESCO Perspective: An OECD-UNESCO Perspective*. OECD Publishing. <https://tinyurl.com/3hbfrxdx>
- Valentine, N., Shanahan, E. M., Durning, S. J., & Schuwirth, L. (2021). Making it fair: Learners' and assessors' perspectives of the attributes of fair judgement. *Medical Education*, 55(9), 1056-066. <https://doi.org/10.1111/medu.14574>
- Wang, Q., Hou, H., & Li, Z. (2022). Participative leadership: A literature review and prospects for future research. *Frontiers in Psychology*, 13, 924357. <https://doi.org/10.3389/fpsyg.2022.924357>
- Weldearegay, C. A. (2020, December). *External quality assurance in Ethiopian higher education: Comparing practices in public and private universities*. (Doctoral thesis-Addis Ababa University). Ethiopia. <https://tinyurl.com/394weya9>
- Wongrugs, A., Kanjanawasee, S., & Ratchusanti, S. (2022). The efficiency of an evaluation model for undergraduate vocational education programs: An application of multiple evaluation approaches. *Kasetsart Journal of Social Sciences*, 43(3), 769-776. <https://doi.org/10.34044/j.kjss.2022.43.3.32>
- Yarbrough, D. B., Shulha, L. M., Hopson, R. K., & Caruthers, F. A. (2010). *The program evaluation standards: A guide for evaluators and evaluation users*. (3rd ed.), Sage Publications. <http://tinyurl.com/39txh4v6>
- Zheng, L., Zhang, X., & Cui, P. (2020). The role of technology-facilitated peer assessment and supporting strategies: A meta-analysis. *Assessment & Evaluation in Higher Education*, 45(3), 372-386. <https://doi.org/10.1080/02602938.2019.1644603>

About the Contributor(s)

Aukkapong Sukkamart, is an Associate Professor in the School of Industrial Education and Technology at the King Mongkut's Institute of Technology Ladkrabang (KMITL) in Bangkok, Thailand. He graduated with a Bachelor of Education in Mathematics and a Master of Education in measurement and research education, and has a Ph.D. in research and curriculum development. He specializes in education research, curriculum development, and measurement/evaluation education.

Email: aukkapong.su@kmitl.ac.th

ORCID: <https://orcid.org/0000-0002-1234-4033>

Jaruwan Ployduangrat, is an Assistant Professor in the School of Industrial Education and Technology at the King Mongkut's Institute of Technology Ladkrabang (KMITL) in Bangkok, Thailand. She graduated with a B.A. in Communication Arts, M.Ed. in Educational Technology, and an Ed.D. in

Educational Administration. She specializes in professional development, teaching and learning, technology-enhanced learning, online education, blended learning, academic writing, and teaching experience.

Email: jaruwan.pl@kmitl.ac.th

ORCID: <https://orcid.org/0009-0005-5870-7740>

Jakkapong Prongprommarat, is an educational supervisor in Thailand's Nakhon Ratchasima Provincial Administration Organizations. He graduated with a Bachelor of Physical Education and a Master of Education in Educational Research and Evaluation. He has a D.Ed. in Educational Administration and a D.I.Ed. in research and curriculum development. He specializes in education research, curriculum development, and measurement and evaluation of education.

Email: pongkued@gmail.com

ORCID: <https://orcid.org/0009-0009-4757-1939>

Paitoon Pimdee, is a Professor with the School of Industrial Education and Technology and Executive Vice President for Procurement, King Mongkut's Institute of Technology Ladkrabang (KMITL), Bangkok, Thailand. He holds a Ph.D. in Environment Education and specializes in educational research, science education, environmental behaviors, conservation behaviors, and statistics.

Email: paitoon.pi@kmitl.ac.th

ORCID: <https://orcid.org/0000-0002-3724-2885>

Publisher's Note: *The opinions, statements, and data presented in all publications are solely those of the individual author(s) and contributors and do not reflect the views of Universitepark, EDUPIJ, and/or the editor(s). Universitepark, the Journal, and/or the editor(s) accept no responsibility for any harm or damage to persons or property arising from the use of ideas, methods, instructions, or products mentioned in the content.*

Appendix I.

Table A1. Population and Sample Distribution of Early Childhood Education (ECE) and Basic Education (BE) Assessors

EQA Supervisory Units (EQASUs)	Population		Sample group		Totals
	ECE	BE	ECE	BE	
1. Progress Education Co., Ltd.	20	45	4	9	13
2. Suranareekanpramurn Co., Ltd	16	44	3	9	12
3. Bangkok Evaluation Co., Ltd	17	41	4	9	13
4. Pattanawittata Limited Partnership	19	44	4	9	13
5. Watchrit Educational Institution Quality Co., Ltd	18	43	4	9	13
6. Sahasikkamatr Ltd., Part	19	41	4	9	13
7. Phionraksa Assessment Limited Partnership	19	42	4	9	13
8. Sueksaphao Co., Ltd	21	43	4	9	13

EQA Supervisory Units (EQASUs)	Population		Sample group		Totals
	ECE	BE	ECE	BE	
9. Surat Assessment Limited Partnership	16	29	3	6	9
10. Chatpetkarnplamern Co., Ltd	16	30	3	7	10
11. K P Assessment Limited Partnership	15	29	3	6	9
12. Narathiwat Assessment Limited Partnership	16	39	3	8	11
13. Wijai Assessment 2003 Limited Partnership	14	45	3	9	12
14. Nakarind Assessment Co., Ltd	14	45	3	9	12
15. U P Quality Assessment Limited Partnership	20	42	4	9	13
16. Thanawat Assessment Limited Partnership	14	42	3	9	12
17. Temprakai Development Limited Partnership	13	42	3	9	12
18. Roi Et Assessment Limited Partnership	15	41	3	9	12
19. Ratchasima Educational Quality Assessment Ltd. Partnership	16	42	3	9	12
20. Khon Kaen Educational Quality Assessment	16	44	3	9	12
21. Mitrplamern Limited Partnership	15	43	3	9	12
22. The Nextgen Evaluation Co., Ltd	16	40	3	8	11
23. Saha Quality Assessment Limited Partnership	15	38	3	8	11
24. Thanaset Kanpramoen Registered Ordinary Partnership	13	39	3	8	11
25. The Best Assessment Co., Ltd	14	38	3	8	11
26. Triple A Educational Quality Evaluation Co., Ltd	15	39	3	8	11
27. PVR Assessment Co., Ltd	14	39	3	8	12
28. Southern Assessment Limited Partnership	14	41	3	9	13
29. Somboon Assess Limited Partnership	18	41	4	9	12
30. Andaman Assessment Co., Ltd	16	44	3	9	12
31. Busarin Assessment Partnership	14	45	3	9	12
32. Rewadee Assessment and Research Partnership	12	43	3	9	12
33. GS Assessment Partnership	12	42	3	9	11
34. Worada Assessment Partnership	14	40	3	8	13

EQA Supervisory Units (EQASUs)	Population		Sample group		Totals
	ECE	BE	ECE	BE	
35. Nonthaburi Assessment Partnership	17	44	4	9	13
36. Arisara Education Quality Assessment Partnership	18	43	4	9	13
37. Songkwae Education Quality Assessment Partnership	17	42	4	9	11
Sub-totals	588	1,514	123	317	440
Population and Sample Totals	2,102		440		

Appendix II.

Table A2. The Questionnaire Components and Their Indicators

Components	Indicators
Organizational Management (OM) (8 indicators)	OM1: The organization has a systematic management structure that emphasizes participation. OM2: The organization can manage its internal operations systematically and emphasize participation. OM3: The organization has a vision and strategies that promote work efficiency in alignment with the policies of ONESQA. OM4: There is an information system related to EQA that is convenient, comprehensive, and secure. OM5: There is a system for supervising and controlling the quality of internal quality assessment reports within the organization. OM6: Coordination between the supervisory unit and external quality assessors is conducted collegially. OM7: EQA planning aligns with the timelines set by ONESQA. OM8: The organization fosters a culture that emphasizes efficiency in EQA.
Personnel performance development (PPD) (5 indicators)	DP1: The organization strengthens the necessary competencies for EQA to meet the needs of the report reviewers and individual external assessors. DP2: External assessors are guided and monitored with the intent that they perform their duties according to the standards, criteria, and methods set by ONESQA with quality. DP3: There is a system for inspecting the EQAs conducted by individual external assessors. DP4: There are processes or methods to support or advise external assessors, enabling them to resolve various issues. DP5: Personnel within the supervisory unit are developed to have the knowledge and capability to review reports systematically.

Components	Indicators
Productivity and results (PAR) (6 indicators)	PAR1: The EQA reports are submitted within the timeline set by ONESQA. PAR2: The EQA reports meet the criteria and principles established by ONESQA. PAR3: The EQA reports and accompanying documents are complete according to the components specified by ONESQA. PAR4: The EQA reports are accepted by educational institutions. PAR5: The EQA reports are valuable and relevant to the context of the educational institutions. PAR6: The satisfaction of stakeholders within the supervisory unit is achieved.
Cooperation Networks (CN) (3 indicators)	CN1: There are effective processes or steps for communication and coordination with ONESQA. CN2: Effective processes or steps for communication and coordination with other organizations exist. CN3: Support is provided for information and data related to EQA that aligns with ONESQA's policies.