

Does Internal Quality Assurance Improve Teaching and Learning? A Systematic Review of Impact Evidence Through the 3P Lens

Abdul Kholid Achmad^{1*}, Sutiah², Rahmat Agus Santoso¹, Wannaporn Siripala³

¹Master of Management Postgraduate Program, Universitas Muhammadiyah Gresik, Indonesia

²Universitas Islam Negeri Maulana Malik Ibrahim, Malang, Indonesia

³Department of Education and Society, Rajamangala University of Technology Krungthep, Thailand

✉ Author Corresponding: abdkholidachmad@umg.ac.id*

ABSTRACT

Internal quality assurance (IQA) has emerged as a central governance mechanism for teaching and learning in higher education, yet its effectiveness remains contested. Prior reviews have primarily clarified how IQA is conceptualized, while empirical evidence on its impact on teaching quality and student learning outcomes remains fragmented and uneven. This systematic review synthesizes evidence of IQA's impact and examines how outcomes vary across macro-meso-micro contexts through the 3P lens (Purpose-People-Process). A comprehensive search was conducted in Scopus, supplemented by backward reference screening. Following PRISMA guidelines, 1,030 records were screened, 96 full-text articles were assessed for eligibility, and 43 studies were retained. Data were extracted using a structured coding framework covering study characteristics, IQA purposes, actors, processes, contextual levels (macro-meso-micro), and reported outcomes. Methodological quality was assessed using the Mixed Methods Appraisal Tool (MMAT). The synthesis combined deductive framework analysis based on the 3P model with realist-informed analysis of context-mechanism-outcome configurations to explain how different IQA arrangements generate, constrain, or moderate impacts on teaching quality and student learning outcomes. Across the included studies, IQA is most consistently associated with strengthened review routines and program-level improvement processes (e.g., curriculum renewal, assurance-of-learning/assessment loops, and governance for monitoring), whereas evidence for direct student learning outcomes is limited and frequently relies on proxy measures such as satisfaction and student evaluations of teaching. Positive impacts are more likely when IQA aligns improvement-oriented purposes with distributed actor engagement and closed-loop processes that translate evidence into action (e.g., feedback-to-action cycles), while compliance-driven configurations tend to show weaker or mixed effects. The review provides an evidence map and a testable program theory linking IQA configurations to outcomes, highlighting priority gaps for longitudinal designs and evaluation research inclusive of the Global South.

Keywords: Internal Quality Assurance; Teaching And Learning Quality; Impact Evidence; 3P Framework, Macro Meso Micro Contexts.



Article History:

Received: 05-05-2026

Revised : 22-06-2026

Accepted: 24-06-2026

Online : 01-08-2026

How to Cite (APA style):

Achmad, A. K., Sutiah, Santoso, R. A., & Siripala, W. (2026). Does Internal Quality Assurance Improve Teaching and Learning? A Systematic Review of Impact Evidence Through the 3P Lens. *IJECA (International Journal of Education and Curriculum Application)*, 9(2), 698-718. <https://doi.org/10.31764/ije.v9i2.39537>



This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license

1. INTRODUCTION

Internal quality assurance (IQA) has emerged as a pivotal governance mechanism in higher education, driven by growing demands for accountability and institutional competition. Effective IQA systems are essential for fostering continuous improvement in teaching and learning, as

evidenced by studies highlighting their role in enhancing the quality of academic programs across various educational contexts, including Vietnam (Pham et al., 2022). Globally, frameworks such as ISO 9001 and guidelines from organizations like INQAAHE and ENQA shape these quality management practices, emphasizing the need for institutions to adopt robust accreditation and audit mechanisms to maintain high academic standards (Miguel et al., 2019; Wysocka, 2022). Furthermore, the implementation of data-driven audits and feedback from stakeholders is crucial for sustaining excellence in private higher education institutions, as these practices enhance institutional trust and competitiveness (Serrano et al., 2025; Sofyani et al., 2023). Overall, the integration of IQA into university governance not only supports institutional accountability but also fosters a culture of continuous quality improvement, benefiting all stakeholders involved in the educational process.

Over the past two decades, internal quality assurance (IQA) has become a near-universal feature of higher education governance. Across Europe, the Bologna Process and the European Standards and Guidelines (ESG) have institutionalised internal quality systems as a core requirement for quality enhancement and accountability. Similar developments have occurred in Asia, Australia, the Middle East, and Latin America, where accreditation agencies increasingly require evidence of systematic internal quality monitoring and continuous improvement. Recent evidence suggests that quality assurance activities now consume substantial institutional resources, including dedicated quality offices, programme review systems, learning outcome assessment mechanisms, and institutional data infrastructures (Hou et al., 2025; Stalmeijer et al., 2023). Despite this widespread adoption, questions remain regarding whether these investments translate into meaningful improvements in teaching quality and student learning outcomes.

The urgency of evaluating IQA effectiveness is further reinforced by persistent concerns regarding the balance between quality improvement and accountability. While quality assurance systems have expanded globally, several studies report that institutions frequently struggle to demonstrate how quality processes contribute to measurable educational outcomes. Instead, quality systems often generate extensive documentation, reporting requirements, and compliance activities without clear evidence of improvements in teaching effectiveness or student achievement (Winstone & Carless, 2021; Mukhatayev et al., 2024). This tension between procedural compliance and educational enhancement has become one of the central debates in contemporary higher education quality management.

Internal Quality Assurance (IQA) in higher education is a dual-edged sword, capable of fostering genuine educational improvements or devolving into a compliance-driven bureaucracy that stifles innovation. While IQA systems are essential for demonstrating accountability and enhancing educational quality, many institutions adopt them merely to satisfy external accreditation requirements, leading to a "tick-box" mentality that prioritizes procedural adherence over substantive quality enhancement (Stalmeijer et al., 2023). Research indicates that a significant number of universities implement isomorphic IQA systems that lack depth and effectiveness, often serving more as a façade for compliance than as a mechanism for genuine improvement (Knyaginina & Puchkov, 2024). Furthermore, faculty engagement is critical for the success of IQA initiatives; however, resistance from faculty can hinder the continuous quality improvement process, emphasizing the need for strategic involvement and recognition to foster a quality culture within institutions (Pham et al., 2022). Thus, while IQA has the potential to drive educational excellence, its effectiveness is contingent upon a commitment to meaningful practices rather than mere compliance (Pramono & Widiyanto, 2024).

Current scholarship on Internal Quality Assurance (IQA) in higher education reveals a fragmented landscape across macro, meso, and micro levels, complicating comparisons of its definitions and functions across various contexts. Research indicates that institutional frameworks and policies significantly influence IQA practices, with established education systems demonstrating more systematic integration of continuous evaluation and professional development compared to those with limited resources (Afiralda et al., 2025; Pujianti et al., 2025). The literature highlights the importance of stakeholder engagement and leadership in fostering a quality culture, emphasizing that effective IQA requires a collaborative approach supported by consistent policy frameworks (Širec & Rožman, 2025). Furthermore, the role of accreditation processes is underscored as essential for maintaining academic standards and enhancing institutional reputation (Iqbal et al., 2025). Overall, the need for a holistic perspective that encompasses diverse dimensions of quality—academic, emotional, and social—is critical for advancing educational outcomes and ensuring accountability.

Although (Krooi et al., 2024) made an important contribution by synthesising 262 studies and introducing the 3P model (Purpose–People–Process) as a conceptual framework for understanding IQA, their review primarily addressed how IQA is configured rather than whether it produces measurable educational impacts. Specifically, the review mapped purposes, actors, and processes associated with IQA implementation but did not systematically evaluate the strength, direction, or mechanisms of impact evidence related to teaching quality and student learning outcomes. Consequently, the field still lacks a synthesis that moves beyond conceptualisation toward explanation of effectiveness. Addressing this gap, the present review narrows its focus to studies reporting outcome-related evidence and uses the 3P framework as an analytical lens to examine how different IQA configurations influence educational outcomes across macro-, meso-, and micro-level contexts. In doing so, this review extends Krooi et al.'s conceptual contribution by developing an evidence-informed programme theory of IQA impact and identifying the contextual conditions under which IQA is most likely to generate meaningful improvements.. This model aligns with the broader educational discourse on collaborative practices and the need for coherent frameworks that integrate various educational resources, as highlighted in studies addressing blended learning and cognitive load management (Lou, 2025). By promoting a contextualized approach, the 3P model facilitates informed decision-making and enhances the overall educational experience, thereby addressing the complexities of quality assurance in diverse learning environments (Freeth & Reeves, 2004).

Empirical evidence on whether internal quality assurance (IQA) improves teaching quality and student learning outcomes remains limited and uneven, with many studies relying on proxy indicators rather than direct measures of learning and instructional change. While recent conceptual work has clarified IQA configurations and their intended functions, it simultaneously underscores that the evidence base for impact on educational quality is still remains limited and methodologically heterogeneous. In practice, institutions often adopt outcome- or evidence-oriented QA architectures (e.g., assurance-of-learning routines and structured outcome assessment) to align goals, enhance monitoring, and support continuous improvement; however, these architectures do not automatically yield robust impact claims unless outcome measures are valid, comparable, and embedded in closed-loop decision processes (Stalmeijer et al., 2023; Vargas et al., 2024). Moreover, accountability and QA agendas can shape what is measured and enacted—privileging what is auditable—so that the “evidence” produced may satisfy procedural requirements but still fall short of demonstrating meaningful improvements in learning (Winstone & Carless, 2021). This challenge is intensified when IQA effectiveness is inferred from

student evaluations of teaching (SETs), whose validity is debated due to systematic bias and fairness concerns; recent large-scale and causal evidence indicates that gendered dynamics can affect evaluation patterns, complicating the use of SETs as a primary indicator of teaching quality (Zheng et al., 2023). Collectively, these strands point to the need for more rigorous, outcome-focused studies that directly test how specific IQA configurations (e.g., improvement-oriented purposes, distributed actor engagement, and feedback-to-action cycles) translate into measurable teaching and learning gains, ideally using longitudinal and quasi-experimental designs and context-sensitive indicators.

Therefore, the central unresolved question is no longer whether higher education institutions have adopted internal quality assurance systems, but whether these systems generate demonstrable educational benefits. By synthesising impact evidence through the 3P lens, this review seeks to explain not only whether IQA works, but also how, for whom, and under what conditions it contributes to improvements in teaching quality and student learning outcomes.

2. METHODS

2.1 Study Design and Reporting Standards

This study adopted a systematic literature review design to synthesize empirical evidence on the impact of Internal Quality Assurance (IQA) in higher education. The review was planned and reported in alignment with PRISMA 2020 principles to ensure transparency and reproducibility across identification, screening, eligibility, and inclusion. The analytical strategy was deliberately oriented toward impact evidence in response to prior conceptual work that highlights the limited and uneven empirical basis for claims about IQA effectiveness in teaching and learning.

2.2 Information Sources and Search Strategy

The literature search was conducted in Scopus on February 2026 because of its broad coverage of peer-reviewed journals in education, management, and policy-relevant fields. Two complementary searches were executed to maximize sensitivity to both (i) core IQA conceptualizations and institutional practices; and (ii) explicit evaluative/impact-oriented studies. The first search used a broad IQA query capturing internal quality assurance, internal quality systems/management, quality enhancement, audit/self-evaluation, and program review language in higher education contexts. The second search added evaluation and outcome terms (e.g., effectiveness, impact, performance, learning outcomes, teaching quality, retention, satisfaction) to preferentially retrieve studies reporting evidence of outcomes attributable to IQA-related processes. The full search strings are reported in the Supplementary Materials to support replication.

To align with the review scope and ensure disciplinary relevance, the Scopus filters were applied at the search stage: publication years 2015–2026, document type “Article”, language “English”, and subject areas Social Sciences, Business/Management, and Decision Sciences. The search was further limited to Open Access records to ensure full-text availability for comprehensive eligibility assessment and extraction.

2.3 Eligibility Criteria

Eligibility criteria were defined a priori to ensure that the included evidence genuinely represented internal quality assurance systems for teaching and learning in higher education. Studies were included when they examined higher education institutions (e.g., universities and colleges), addressed IQA as an institutional or program-level system (e.g., internal audits, self-

evaluation, program review, PDCA/continuous improvement, quality enhancement infrastructures), and either described IQA governance and implementation (purpose–people–process elements) or evaluated IQA processes and outcomes. Empirical studies using qualitative, quantitative, or mixed-methods designs were eligible, and conceptual/framework papers were also eligible when they provided definitional clarity or model-building directly relevant to operationalizing IQA in higher education.

Studies were excluded when they focused on non-higher education settings, addressed external quality assurance/accreditation without an internal system or process, used “quality” terminology in a generic sense without QA/IQA content, or were situated in non-education/industry contexts without clear relevance to IQA in higher education. Full-text inaccessibility at the eligibility stage was treated as a basis for exclusion only when content was insufficient to judge relevance after reasonable retrieval efforts, with the reason documented.

The literature search was conducted in Scopus because of its broad coverage of peer-reviewed higher education and quality assurance research. The search strategy was developed iteratively through preliminary scoping searches and refinement of key concepts related to internal quality assurance (IQA) and higher education. The final search string combined terms representing internal quality assurance systems, quality management processes, and quality enhancement mechanisms with terms representing higher education institutions using Boolean operators (AND/OR). Truncation symbols and phrase searching were applied where appropriate to maximise sensitivity while maintaining relevance.

The final search query was: TITLE-ABS-KEY(("internal quality assurance" OR "institutional quality assurance" OR "academic quality assurance" OR "internal quality management" OR "internal quality system" OR "internal quality systems" OR "quality assurance system" OR "quality assurance systems" OR "quality assurance framework" OR "quality assurance frameworks" OR "quality assurance mechanism" OR "quality assurance mechanisms" OR "quality assurance process" OR "quality assurance processes" OR "quality enhancement" OR "quality management system" OR "quality management systems") AND ("higher education" OR universit* OR college* OR "tertiary education" OR "postsecondary education" OR "higher education institution" OR "higher education institutions" OR HEI))

Additional filters were applied to retain only journal articles published in English between 2015 and 2026 within the Social Sciences, Business/Management, and Decision Sciences subject areas. Open-access availability was not used as an eligibility criterion but facilitated full-text retrieval during screening. The complete search strategy, screening decisions, and exclusion criteria are provided in Figure 1.

2.4 Study Selection and Screening Process

All records were exported from Scopus in RIS format and consolidated into a single screening database. Deduplication was performed using DOI as the primary key, followed by secondary checks on title and source to minimize false duplicates. The combined searches yielded 1,067 records, of which 37 were duplicates, resulting in 1,030 unique records for screening. Screening proceeded in two stages. First, titles and abstracts were screened against the eligibility criteria to remove clearly irrelevant studies while retaining borderline cases for full-text inspection. To improve consistency, eligibility rules were calibrated iteratively during the early stages of screening, and ambiguous cases were retained for full-text assessment rather than excluded prematurely. Second, full texts were assessed for eligibility, and each exclusion at the full-text stage was documented with an explicit reason code aligned to the predefined criteria. Following

title/abstract screening, 96 articles proceeded to full-text assessment; 53 were excluded at full text for not meeting the operational definition of IQA in higher education or for lacking sufficient relevance to internal QA systems for teaching and learning, leaving 43 studies for final inclusion. A PRISMA flow diagram summarizes the selection process and counts across stages (Figure 1).

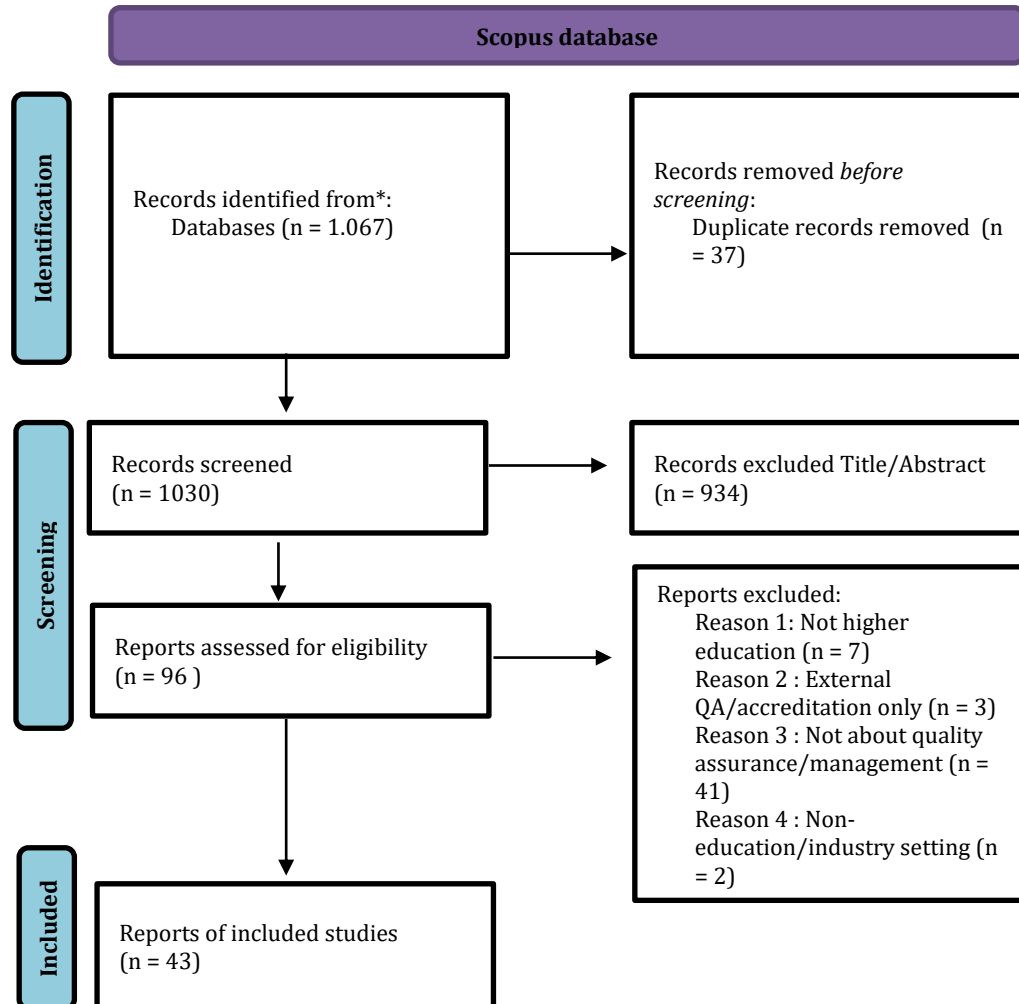


Figure 1. PRISMA flow diagram of identified and included studies

2.5 Data Extraction and Coding Framework

A structured extraction template was developed to ensure consistency and transparency throughout the review process. The template was pilot-tested on a subset of included studies and refined iteratively before full extraction commenced. For each study, descriptive information was extracted, including author(s), publication year, country or region, higher education context, study design, sample characteristics, and data collection methods. To support the impact-focused objectives of the review, additional analytical variables were extracted, including reported teaching-quality outcomes, student-learning outcomes, outcome indicators, and evidence sources.

The coding framework was informed by the 3P model of internal quality assurance (Purpose–People–Process) proposed by Krooi et al. (2024). Accordingly, studies were coded according to their dominant IQA purpose (e.g., improvement, accountability, compliance, or hybrid), principal actors involved in implementation (e.g., leadership, faculty, quality assurance units, students, external stakeholders), and key quality assurance processes (e.g., programme review, assessment

of learning outcomes, feedback systems, continuous improvement cycles). To capture contextual variation, studies were also coded according to macro–meso–micro analytical levels. Macro-level codes represented policy, regulatory, accreditation, or system-level influences; meso-level codes represented institutional governance and organisational quality structures; and micro-level codes represented programme, classroom, faculty, or learner-level implementation processes. Finally, outcome evidence was coded according to a three-stage impact pathway comprising outputs, intermediate outcomes, and end outcomes. This structure facilitated comparison across studies and supported the subsequent framework synthesis and realist-informed analysis of context–mechanism–outcome relationships.

2.6 Quality Appraisal and Evidence Strength

Methodological quality was assessed using the Mixed Methods Appraisal Tool (MMAT, 2018), which enables the appraisal of qualitative, quantitative, and mixed-methods studies within a single review framework (Hong et al., 2018). Depending on study design, each article was evaluated against five methodological criteria addressing the appropriateness of research questions, sampling strategy, data collection procedures, measurement validity, risk of bias, and coherence between data sources and interpretations. Quality assessment was conducted independently by the review team using a standardised appraisal form derived from the MMAT manual. Disagreements were resolved through discussion and consensus. Following MMAT guidance, individual criterion ratings were reported rather than aggregated into a single numerical quality score.

To support interpretation of review findings, methodological quality assessments were complemented by an evidence-strength classification adapted from evidence synthesis literature. Studies were categorised as strong, moderate, or limited evidence based on methodological quality, outcome measurement validity, and the extent to which causal inferences regarding IQA impact could be supported. Studies employing longitudinal, mixed-methods, or multiple-source outcome measures were generally considered stronger evidence than studies relying solely on cross-sectional perceptions or proxy indicators. These classifications informed interpretation of findings but were not used as exclusion criteria.

2.7 Data Synthesis and Analysis

Given the heterogeneity of contexts, study designs, and outcome measures, meta-analysis was not considered appropriate. Instead, an integrative synthesis approach was employed. First, an evidence map was developed to describe the distribution of impact outcomes across macro–meso–micro contexts and across regions, including Global North/Global South categorization, to address documented geographical imbalances in IQA research. Second, impact findings were synthesized using a structured narrative approach supported by “direction of effect” summarization, distinguishing consistent positive effects, mixed/contingent effects, and null/unclear effects, while explicitly differentiating direct learning/teaching outcomes from proxy indicators such as satisfaction or perceived service quality. Third, a framework synthesis was conducted using the 3P model to connect purposes, actors, and processes to observed outcomes, enabling the review to move beyond whether IQA “works” to explain how and under what conditions impact is more likely to emerge. Finally, mechanisms and contextual conditions were integrated into an interpretive account of plausible causal pathways from IQA processes to outcomes, highlighting testable propositions and priority gaps for future longitudinal and context-sensitive evaluation studies.

3. RESULT AND DISCUSSION

3.1 Study Selection

The database search retrieved 1,067 records, of which 37 duplicates were removed, leaving 1,030 unique records for title/abstract screening. Following title/abstract screening, 96 articles were retained for full-text assessment. Full-text screening excluded 53 articles for not meeting the operational definition of IQA in higher education teaching and learning, yielding 43 studies for final inclusion (Figure 1).

3.2 Characteristics of Included Studies

The evidence base increased over time, with the largest concentration in 2025 ($n = 9$) and sustained output from 2020–2025 (Figure 2a). Geographically, the literature was dominated by Global North settings (74.4%), with Global South evidence comprising 20.9% and a small cross-context subset (4.7%). Methodologically, qualitative and mixed-methods designs dominated, while purely quantitative designs constituted approximately one-fifth of included studies. In terms of contextual level, nearly half of studies focused on meso–micro enactment (institutional arrangements linked to program/faculty implementation), whereas macro–meso analyses were less common and macro–micro linkages were rare (Figure 2b).

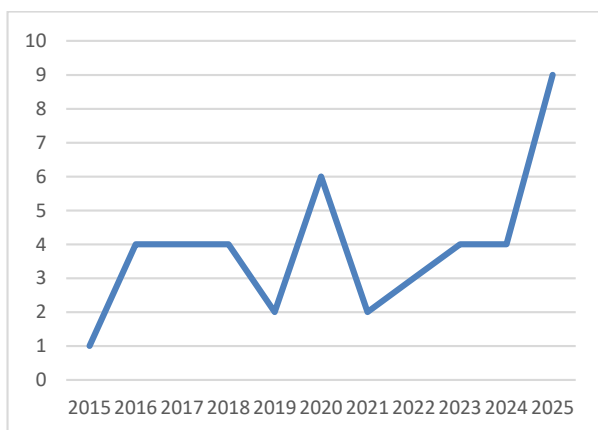


Figure 2a. Publication year distribution

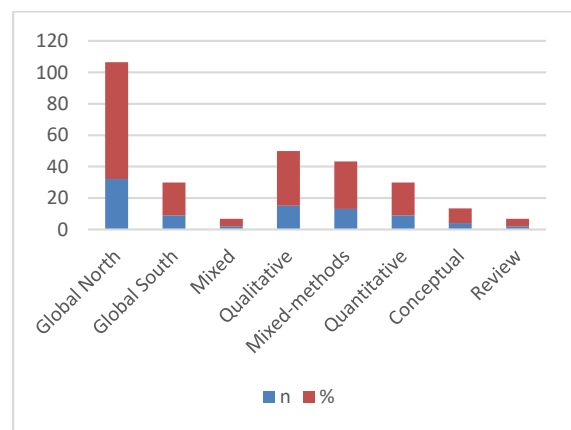


Figure 2b. Summary characteristics

3.3 How IQA is Configured for Teaching and Learning (3P Mapping)

Purpose (why IQA is enacted)

Across the included studies, IQA was most frequently framed as improvement-oriented ($n = 19$), followed by hybrid purposes that combine improvement and accountability logics ($n = 11$). A smaller portion positioned IQA primarily as compliance/accountability ($n = 5$), suggesting that much of the literature normatively prioritizes enhancement but frequently operates under mixed governance rationales (Table 1).

Table 1. Purpose Orientation

Purpose orientation	n	%
Improvement-oriented	19	44.2
Hybrid (improvement + accountability)	11	25.6
Other/unclear	8	18.6
Accountability/compliance-oriented	5	11.6

People (who drives and enacts IQA)

The actor landscape was multi-stakeholder, with faculty/academic staff ($n = 29$) and students ($n = 25$) commonly appearing alongside external stakeholders ($n = 26$; e.g., employers, agencies, accreditation-related interfaces). Leadership/management was explicitly central in fewer studies ($n = 12$), while dedicated QA units/“third space” professionals were explicitly foregrounded in a small subset ($n = 3$), indicating that QA work is often described through program-level and stakeholder-facing practices rather than through formal QA office structures alone (Table 2).

Table 2. Actor Groups Mentioned in IQA Enactment

Actor group	n	%
Faculty/academic staff	29	67.4
External stakeholders (employers/agency)	26	60.5
Students	25	58.1
Leadership/management	12	27.9
QA unit / third space professionals	3	7.0

Process (what IQA does in practice)

Process descriptions clustered into seven recurring instrument families, with learning-outcomes assessment (including assurance-of-learning loops), student feedback systems, program review/curriculum reform cycles, and governance infrastructures appearing most often. These process families are analytically important because they constitute the “intervention layer” through which impact is expected to materialize (Table 3).

Table 3. IQA Process/Instrument Families

Process / instrument family	n	%
Learning-outcomes assessment / AoL loops	13	30.2
Student feedback & evaluation systems (SET/SSS)	9	20.9
Program/course review & curriculum reform cycles	8	18.6
Governance structures (QA units, committees, bylaws, policy)	8	18.6
Continuous improvement routines (PDCA/closing-the-loop)	5	11.6
Performance indicators, KPI dashboards, benchmarking	4	9.3
Self-evaluation & internal audit	3	7.0

3.4 What “Impact Evidence” Exists (Measured End Outcomes VS Process-Only Evidence)

To align with an impact-evidence emphasis, the included studies were separated into those reporting measured end outcomes (impact evidence) versus those primarily reporting process/system outputs or conceptual contributions without measurable outcomes. Using this operational definition, 22 of 43 studies (51.2%) provided measured end outcomes. Within this subset, the most common outcome domains were institutional/distal outcomes (27.3%), teaching/curriculum outcomes (22.7%), and direct learning outcomes (18.2%), with the remainder relying on proxy/perceptual indicators or mixed teaching-proxy outcomes (Table 6). Impact-evidence studies were spread across qualitative, quantitative, and mixed-methods designs, but remained predominantly situated in Global North contexts (68.2%), with Global South evidence forming 27.3% (Table 4).

Table 4. Impact-Evidence Subset Profile (n = 22)

Characteristic	Category	n	%
Outcome category (impact subset)	Institutional/distal outcomes	6	27.3
	Teaching/curriculum outcomes	5	22.7
	Direct learning outcomes	4	18.2
	Teaching & proxy outcomes	4	18.2
	Proxy/perceptual outcomes	3	13.6
Design (impact subset)	Qualitative	8	36.4
	Quantitative	7	31.8
	Mixed-methods	7	31.8
Region (impact subset)	Global North	15	68.2
	Global South	6	27.3
	Mixed	1	4.5

Direct learning outcomes constituted a smaller share of the impact subset than teaching/curriculum or institutional outcomes. The MMAT appraisal indicates that the confidence of conclusions varies by outcome domain and study design. Overall, the evidence base is dominated by cross-sectional and perception-based studies (e.g., student satisfaction and student evaluations of teaching), which often meet basic reporting and analysis requirements but provide limited leverage for causal attribution. As a result, conclusions about IQA's association with strengthened review routines, governance, and program-level improvement processes can be stated with moderate confidence, because these outcomes are frequently documented through triangulated qualitative evidence and mixed-method evaluations that describe implementation processes and organizational change. In contrast, conclusions regarding IQA's impact on direct student learning outcomes remain low-to-moderate confidence, because fewer studies use validated learning outcome measures, longitudinal/pre-post designs, or credible counterfactual logic; many rely on proxies that are sensitive to context and may not reflect learning gains. Methodological quality therefore shapes the interpretation of "impact" in this review: stronger studies support claims that IQA improves institutional improvement infrastructures and closing-the-loop capacity, while weaker or indirect evidence means that claims about learning gains should be treated as tentative and contingent on configuration, context, and measurement strategy. Consistent with MMAT guidance, quality appraisal was used to qualify interpretive weight rather than to mechanically exclude studies; the review thus highlights where evidence is robust, where it is suggestive, and where future research requires more rigorous designs and outcome measures.

3.5 Evidence Maps Across Contextual Levels (All Included Studies)

Across the full set of included studies, outcome reporting varied by context level. Meso-micro studies were the dominant location for teaching/curriculum outcomes and direct learning outcomes, whereas macro-meso studies more often reported institutional/distal outcomes and system/process outcomes. Figure 3 presents the evidence map linking outcome categories to contextual level (macro-meso-micro).

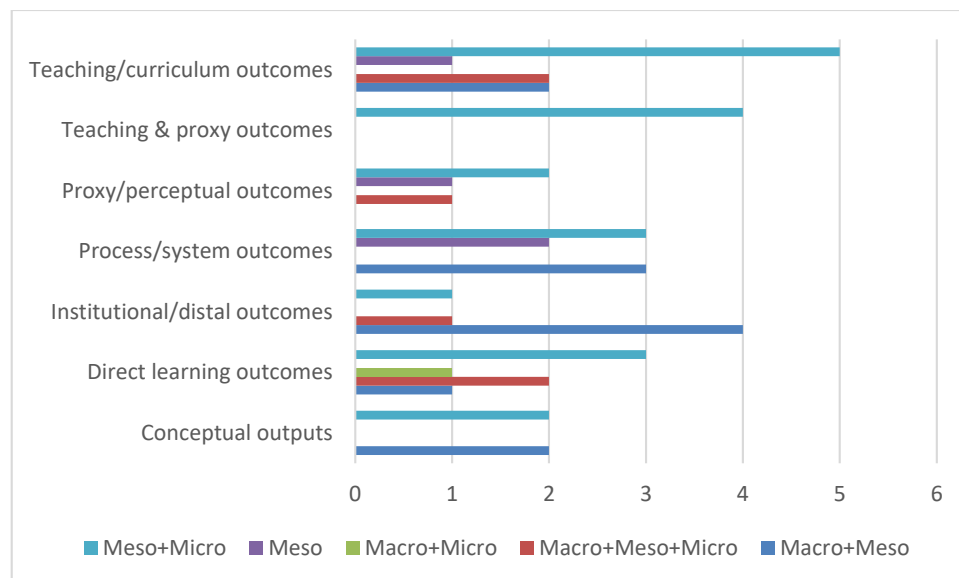


Figure 3. Evidence map (outcome category × contextual level)

3.6 Mechanisms And Conditions Associated With Iqa-Related Outcomes (Reported Themes)

The synthesis of extracted mechanism descriptions indicated that IQA impact narratives most frequently rested on closed-loop improvement logics (e.g., PDCA and feedback-to-action cycles) and organizational learning/quality culture framing. Participatory governance and stakeholder engagement mechanisms were also prominent, whereas explicit accreditation/rankings pressure mechanisms were reported less often in the included set (Table 5).

Table 5. Mechanism themes reported in included studies (n = 43)

Mechanism theme	n	%
Closed-loop improvement (PDCA/feedback-to-action)	18	41.9
Organizational learning / quality culture	17	39.5
Stakeholder/participatory governance	9	20.9
Performance management / KPI logics	6	14.0
Intellectual/social capital lens	4	9.3
Accreditation/rankings pressure	2	4.7

In terms of contextual conditions, stakeholder engagement and data usability were the most frequently reported enabling conditions, while barriers were more concentrated around resource/time constraints and bureaucratic/compliance burdens conditions that potentially attenuate translation from IQA routines to observable educational outcomes (Table 6).

Table 6. Reported enabling conditions and barriers

Theme	Enablers (n)	Barriers (n)
Stakeholder engagement	20	6
Data systems & usability	8	6
Quality culture / shared ownership	7	1
Leadership support	6	2
Resistance / trust issues	5	5
Capacity & professional development	3	3
Resources & time	2	14
Compliance burden / bureaucracy	2	10

3.7 Discussion

Beyond identifying whether IQA is associated with positive, mixed, or limited outcomes, the review sought to explain why impact varies across studies. The synthesis indicates that variation is best understood through differences in Purpose–People–Process configurations rather than through the mere presence or absence of quality assurance systems. Consequently, the discussion focuses on how different IQA configurations generate distinct outcome patterns and what these patterns imply for theory and practice. Our theoretical contribution is not another typology of IQA tools, but an explanatory model that specifies the mechanism chain and the boundary conditions under which IQA configurations translate into educational outcomes.

a. Positioning and Contributions: From 3P Conceptualization to Impact Evidence

This review advances the 3P agenda introduced by (Krooi et al., 2024) by shifting the focus from conceptual clarification toward impact evidence, namely what internal quality assurance (IQA) demonstrably delivers for teaching and learning outcomes in higher education and under what conditions such impacts are most likely. While the 3P model offers a shared language to describe IQA configurations (Purpose–People–Process), the central gap identified by Krooi et al. (2024) a limited and uneven evidence base for IQA's impact requires a synthesis that explicitly distinguishes process outputs, intermediate outcomes, and end outcomes, and that is sensitive to contextual variation across macro–meso–micro levels.

The results show a growing but still uneven field out of 43 included studies, only 22 reported end-outcome evidences (impact subset), with substantial heterogeneity in outcome definitions, measurement strategies, and evaluation designs. This pattern aligns with broader quality assurance scholarship arguing that QA often becomes most visible through routines, artefacts, and compliance infrastructures, while educational impact depends on whether QA information is translated into action in the classroom and program space (Winstone & Carless, 2021). Thus, the main contribution of this review is not simply to report whether IQA “works,” but to explain why impact evidence is thin, where it is strongest, and how alignment across purpose–people–process conditions the likelihood of measurable improvement.

b. What IQA Impacts Most Consistently: Stronger Signals In System Routines Than Learning Gains

The evidence in this review indicates that IQA is most consistently associated with improvements in institutional and program routines—such as review cycles, evaluation practices, curriculum redesign processes, and strengthened governance for monitoring—rather than with direct evidence of student learning gains at scale. In several studies, IQA was credibly linked to strengthened review infrastructures or governance and stakeholder engagement processes that make teaching quality more discussable and auditable (Gray et al., 2022; Hoare et al., 2024). These forms of improvement are meaningful; however, they do not necessarily establish that IQA produces measurable learning impact, particularly when the measurement system relies on proxies such as perceived service quality or satisfaction (Rahman & Nasrin, 2024).

A second defining feature of the evidence base is the prevalence of proxy outcomes, including perceptions of quality, stakeholder satisfaction, and self-reported effectiveness. For example, work connecting quality management practices to stakeholder satisfaction and studies examining perceived service quality (Rahman & Nasrin, 2024) are valuable for understanding stakeholders' interpretations of “quality,” but their relationship to

learning outcomes is indirect. This pattern is consistent with concerns that QA can become “measurement-heavy” yet pedagogically “impact-light” unless feedback is designed to support improvement and sensemaking (Winstone & Carless, 2021). Consequently, the evidence base currently supports a cautious inference: IQA more reliably improves quality-related routines and organizational learning conditions than it demonstrates direct learning-outcome gains, which empirically reinforces the gap emphasized (Krooi et al., 2024).

c. Explaining Impact Through the 3P Lens: Alignment is the Boundary Condition

A key insight from mapping studies to the 3P dimensions is that the likelihood of demonstrable impact depends less on the presence of QA procedures per se than on the alignment of purpose, people, and process. Where IQA is framed primarily as improvement-oriented or explicitly hybrid (improvement + accountability), the processes described are more likely to form closed-loop routines that support decision-making and pedagogical change (Krooi et al., 2024). Conversely, where IQA is framed as compliance-centric, activities tend to emphasize documentation, reporting, and procedural completion, which may increase transparency but does not necessarily trigger teaching-and-learning improvements unless complementary enabling conditions exist.

The “People” dimension provides further explanatory power: studies in this review repeatedly suggest that IQA produces stronger outcomes when it is not confined to specialized quality offices but becomes a distributed organizational project supported by leadership, engaged faculty, and credible stakeholder roles (Mazza & Azzali, 2025; Širec & Rožman, 2025). This resonates with empirical findings that stakeholder involvement and governance structures affect how QA standards are interpreted and operationalized at program level, particularly in audit contexts. It also aligns with the argument that quality work requires “quality literacy,” i.e., stakeholders’ competence to interpret QA frameworks, understand trade-offs, and integrate quality practices into day-to-day academic work (Leiber & Seyfried, 2025). Without such interpretive capacity, IQA risks producing compliance artefacts rather than instructional improvement.

Finally, the “Process” dimension indicates that impact is most plausible when IQA processes operate as closed-loop systems—for example, program review and assurance-of-learning cycles that connect evidence to action and re-measurement. Studies proposing structured learning outcome frameworks and assessment loops underscore the importance of making outcomes explicit and auditable in ways that support improvement rather than merely satisfy external demands (Vargas et al., 2024). Similarly, empirical evidence on internal QA units suggests that organizational structures can support or hinder impact depending on whether they are integrated into institutional strategy and decision processes rather than operating in parallel to academic work (Erittu & Turri, 2023). Taken together, the 3P synthesis implies that alignment is the operative boundary condition: misalignment between improvement aspirations, actor engagement, and process design predicts weaker or mixed outcomes even when IQA systems formally exist.

d. Macro, Meso, Micro Contingencies and the Global North/South Imbalance

A central challenge for impact-oriented conclusions is that IQA does not operate in a vacuum; it is shaped by macro-level policy regimes, meso-level governance arrangements, and micro-level enactment. Evidence map (Fig.7) shows that meso-micro studies are the dominant location where teaching- and program-level outcomes appear, while macro-meso studies more often report system outputs and institutional performance outcomes.

This pattern suggests that macro-level pressures (e.g., accreditation, rankings, national QA regimes) can increase accountability but do not automatically translate into learning improvement unless meso-level governance is designed to enable micro-level pedagogical change (Hou et al., 2025). In other words, accountability may intensify IQA activity, but impact depends on whether quality routines reshape teaching practices and curriculum decisions in concrete ways.

The geographical distribution of evidence further limits generalizability. The included literature is dominated by Global North settings, with fewer studies in Global South contexts. This is not merely a representational issue; contexts differ in governance capacity, resource constraints, and the institutional maturity of QA infrastructures. Stakeholder-based evidence from Kazakhstan, for example, shows that internal QA systems can become formalized without being decision-relevant, and that quality services may be perceived as “parallel” or disciplinary rather than developmental units unless data are actually used in institutional strategy and practice (Mukhatayev et al., 2024). In Global South settings, where resource and capacity constraints are often more visible, the translation from QA routines to pedagogical improvement may require stronger enabling conditions—especially capacity-building for staff and stakeholder engagement mechanisms that are not tokenistic.

Importantly, several studies point to IQA as a socially constructed practice: how academics and stakeholders interpret QA agendas influences whether IQA becomes developmental or punitive. For example, the enactment of feedback processes under QA and accountability agendas can shift “for whom” feedback is produced, potentially weakening its pedagogical function if it becomes primarily performative or compliance-driven (Winstone & Carless, 2021). Thus, impact evidence should be interpreted as context-dependent: IQA may “work” differently depending on the macro governance logic and the organizational culture that mediates implementation.

e. Mechanisms and Conditions: What Enables (or Blocks) IQA from Producing Impact

Mechanism synthesis across the included studies suggests that IQA impact narratives most frequently rest on organizational learning and closed-loop improvement logics (e.g., PDCA-like cycles, feedback-to-action loops, or repeated program review with follow-through). These mechanisms are more explicit in studies where IQA is described as a learning system rather than as an audit system. In practice, this typically requires processes that are not only evaluative but also actionable, including clear decision rights, resourcing of improvement initiatives, and re-measurement strategies (Krooi et al., 2024; Vargas et al., 2024).

The enabling and constraining conditions observed in your Results reinforce this mechanism account. Stakeholder engagement and data usability appear frequently as enablers, implying that IQA is more likely to yield outcomes when evidence is credible, timely, and interpretable by those responsible for pedagogical decisions (Leiber & Seyfried, 2025; Mukhatayev et al., 2024). Conversely, resource/time constraints and compliance burdens appear disproportionately among barriers, suggesting that many IQA systems struggle not with collecting data but with converting data into sustained improvement actions. This reading aligns with the broader literature on quality management in higher education showing that stakeholder satisfaction and perceived quality are sensitive to the degree to which quality management practices are coordinated and meaningful rather than fragmented.

A particularly instructive strand of the evidence base concerns program-level QA and assessment. Studies using course-based program review, bylaws analysis, and competency-based learning outcomes frameworks illustrate how “impact evidence” depends on governance for action: program review matters when it is embedded in decision-making and followed by concrete change and re-evaluation (Gray et al., 2022; Hoare et al., 2024; Vargas et al., 2024). In short, the recurring weak link is often implementation fidelity of the improvement loop, not the presence of QA policies.

f. Linking Purpose–People–Process Configurations to Outcomes

A key finding of this review is that IQA outcomes are not determined by the presence of quality assurance systems alone but by how Purpose, People, and Process are configured and aligned. Across the included studies, positive outcomes were most frequently reported when improvement-oriented purposes were combined with broad stakeholder engagement and closed-loop quality processes. In these configurations, IQA functioned as an organisational learning mechanism that translated evaluative information into curriculum enhancement, pedagogical improvement, and evidence-informed decision-making.

In contrast, studies reporting limited or mixed impacts were commonly characterised by compliance-dominated purposes, narrow ownership concentrated within quality assurance offices, and process structures focused primarily on documentation and reporting. Although such configurations often generated procedural outputs (e.g., reports, audits, accreditation readiness), they rarely demonstrated clear effects on teaching quality or student learning outcomes. The synthesis therefore suggests that IQA effectiveness depends less on the existence of quality assurance mechanisms and more on the alignment between improvement-oriented purposes, distributed actor engagement, and actionable quality processes. This finding extends previous conceptual work by demonstrating that different 3P configurations are associated with different outcome trajectories.

g. Theoretical Contribution: a 3P-Based Program Theory of IQA Impact

This review formulates a 3P-based program theory that explains heterogeneity in IQA impact evidence by specifying how Purpose–People–Process configurations operate across macro–meso–micro contexts to produce (or fail to produce) educational outcomes. At the macro level, accountability and competitive pressures can intensify IQA activity and prioritise auditable evidence, but they do not ensure learning improvement unless meso-level governance enables micro-level pedagogical change (Hou et al., 2025; Winstone & Carless, 2021). At the meso level, the decision relevance of IQA depends on whether quality structures are integrated into institutional strategy, resourcing, and decision routines, rather than functioning as parallel administrative systems (Erittu & Turri, 2023). At the micro level, educational impact depends on enactment: whether academics and stakeholders possess sufficient quality literacy to interpret evidence and implement improvements, and whether improvement cycles are closed through action and re-measurement (Leiber & Seyfried, 2025; Stalmeijer et al., 2023).

The program theory specifies a three-stage causal chain. First, IQA commonly produces outputs (policies, reports, indicators, audit artefacts). Second, when alignment and capacity exist, these outputs generate intermediate outcomes (stronger program review routines, improved governance for monitoring, shared quality culture). Third, end outcomes (improved teaching practice and student learning outcomes) emerge only when IQA is configured as a closed-loop system that (i) uses valid outcome measures, (ii)

converts evidence into actionable decisions, and (iii) sustains follow-through via re-measurement and pedagogical support (Benson et al., 2024). Consistent with the MMAT appraisal rationale, the review's confidence is therefore higher for process/system outcomes and lower for direct learning outcomes, given the predominance of cross-sectional designs and proxy indicators in the existing literature (Krooi et al., 2024; Winstone & Carless, 2021).

h. Formal Propositions Derived from the Synthesis

Drawing on the integrated findings, the synthesis supports a set of formal propositions that specify how IQA configurations are expected to relate to outcomes and why effects vary across contexts. First, the review suggests that purpose orientation is consequential: IQA systems enacted primarily for improvement or hybrid systems that retain explicit safeguards for improvement are more likely to be associated with teaching-quality and curriculum outcomes than IQA enacted predominantly for compliance and accountability. This proposition follows from the recurrent pattern that improvement-oriented purposes activate developmental logics (learning, experimentation, refinement), whereas compliance-oriented purposes more often privilege documentation and procedural completion, which may increase visibility but does not reliably translate into educational impact (Krooi et al., 2024; Winstone & Carless, 2021).

Second, the synthesis indicates that IQA impact depends on who owns and enacts the system. When ownership is distributed across leadership, faculty, and quality professionals, the evidence-to-action pathway is strengthened because interpretive capacity and legitimacy are higher and follow-through is more feasible. Conversely, when IQA work is largely delegated to administrative quality functions with limited academic ownership, translation into teaching practice is weaker and outcomes are more likely to be mixed or symbolic. This proposition is consistent with evidence that internal QA units can enable or constrain improvement depending on their integration into institutional decision routines and the extent to which academic actors regard IQA as relevant to their core work (Erirtu & Turri, 2023; Leiber & Seyfried, 2025).

Third, the synthesis supports a proposition about process design: IQA yields more consistent program-level improvement when it is operationalized as a closed-loop system—feedback-to-action with decision rights, resourcing, and re-measurement—than when it is dominated by audit and reporting routines without closure. In other words, the mechanism of impact is not “evaluation” per se but the capacity to convert evaluative evidence into decisions and implemented changes that are subsequently re-assessed. This proposition aligns with scholarship on quality culture and continuous enhancement, as well as with learning-outcome assessment approaches that emphasize actionable measurement and iterative improvement cycles (Stalmeijer et al., 2023; Vargas et al., 2024).

Fourth, the synthesis highlights measurement validity as a boundary condition for detecting impact on student learning. Observable effects on learning outcomes are more likely when institutions adopt valid and comparable learning outcome measures rather than relying mainly on proxy indicators such as satisfaction or student evaluations of teaching (SET) alone. Proxy-heavy measurement reduces the capacity to detect learning gains and increases vulnerability to bias and confounding, which helps explain why the evidence base is stronger for process and routine improvements than for direct learning outcomes. This proposition is consistent with concerns that QA agendas can reshape

feedback practices toward accountability demands and with calls for pragmatic outcome frameworks that maintain validity while remaining feasible in institutional settings (Vargas et al., 2024; Winstone & Carless, 2021).

Fifth, the synthesis points to a macro-meso-micro proposition: macro-level accountability pressures (e.g., accreditation and competitive regimes) may intensify IQA activity, but positive educational impacts occur primarily when meso-level governance enables micro-level enactment through actionable decision routines, time allocation, and resourcing. Where macro pressure is not matched by meso-level enablement and micro-level capacity, IQA activity is more likely to generate compliance outputs than improvements in teaching and learning. This proposition is supported by evidence showing how external drivers shape institutional behavior and by the multi-level configuration logic embedded in the 3P framework (Hou et al., 2025; Krooi et al., 2024).

Sixth, the synthesis supports a moderation proposition regarding quality literacy. The effect of IQA on improvement outcomes is strengthened when stakeholders possess the competence to interpret quality evidence and to design improvement actions, whereas limited quality literacy encourages procedural compliance and weakens improvements in teaching practice. This proposition helps explain why similar IQA structures can yield different outcomes across institutions and systems, and it highlights the importance of interpretive capacity and institutional capability, not only structural arrangements (Leiber & Seyfried, 2025).

Seventh, the synthesis supports a mediational proposition concerning professional development. The relationship between IQA feedback and improved teaching practice is more likely to materialize when professional development and pedagogical support translate QA signals into instructional capability; without such supports, IQA tends to strengthen monitoring routines more than it changes teaching practice. This proposition is consistent with evidence from blended-learning QA interventions, which indicates that multi-component strategies that include professional learning and communities of practice are more promising than isolated compliance mechanisms (Benson et al., 2024). Finally, the synthesis yields a contextual contingency proposition relevant to Global South settings: in resource-constrained environments, IQA impact is more sensitive to capacity and data usability, and without these enabling conditions IQA is more likely to become formalized but not decision-relevant, thereby attenuating effects on teaching and learning outcomes. Stakeholder-based accounts of QA challenges illustrate how limited capacity can produce parallel systems that meet formal requirements yet fail to drive improvement decisions, underscoring the need for context-sensitive evaluation and capacity-building in underrepresented systems (Mukhatayev et al., 2024).

i. Extending the 3P Model: From Configuration to Impact

The findings extend Krooi et al. (2024) 3P model by demonstrating that Purpose, People, and Process operate not merely as descriptive dimensions of IQA systems but as interacting mechanisms influencing educational outcomes. While the original 3P model clarified how internal quality assurance systems are configured, the present review provides evidence regarding how these configurations shape impact.

Specifically, the synthesis suggests that Purpose establishes the dominant institutional logic of quality assurance, determining whether IQA is enacted primarily for improvement or compliance. People influence the translation of quality information into action through ownership, participation, and quality literacy. Process determines whether evaluative

information remains symbolic or is transformed into sustained improvement through feedback and follow-up mechanisms.

Viewed through this lens, IQA can be conceptualised as an organisational learning system. Educational impact emerges when information generated through quality assurance processes is interpreted, acted upon, and institutionalised through repeated improvement cycles. The review therefore extends the 3P framework from a classification model to an explanatory framework capable of accounting for variation in observed outcomes across higher education contexts, as shown in Table 7.

Table 7. Outcome Patterns Across 3P Configurations

Dominant Purpose	Dominant People Configuration	Dominant Process Configuration	Typical Outcomes
Improvement	Distributed ownership	Closed-loop improvement cycle	Strong positive
Improvement	Limited ownership	Partial follow-up	Moderate positive
Compliance	QA-office dominated	Audit/reporting	Weak
Compliance	Administrative ownership	Documentation-focused	Mixed
Hybrid	Shared ownership	Feedback-action cycle	Positive but variable

4. CONCLUSION

This systematic literature review examined the impact of Internal Quality Assurance (IQA) on teaching quality and student learning outcomes in higher education through the analytical lens of the Purpose–People–Process (3P) framework. Synthesising evidence from 43 studies, the review demonstrates that IQA is more consistently associated with improvements in organisational processes, programme review practices, quality culture, and evidence-informed decision-making than with direct improvements in student learning outcomes. Evidence of learning impact remains limited and is frequently based on proxy indicators such as student satisfaction or course evaluations rather than robust measures of learning achievement.

Beyond documenting outcome patterns, the review contributes theoretically by extending 3P framework from a descriptive model of IQA configuration to an explanatory model of IQA impact. The findings suggest that educational outcomes are not produced by the existence of quality assurance systems alone but by the alignment of improvement-oriented purposes, distributed stakeholder engagement, and closed-loop quality processes. In this regard, the review advances a programme theory in which Purpose establishes the dominant logic of quality assurance, People shape the translation of quality evidence into action, and Process determines whether improvement cycles are sustained over time. This theoretical extension helps explain why similar IQA systems generate different outcomes across institutions and contexts.

The findings also offer practical implications for higher education leaders, quality managers, and policy makers. First, institutions should move beyond compliance-oriented quality assurance and explicitly position IQA as a continuous improvement mechanism linked to teaching and learning enhancement. Second, ownership of quality assurance should be distributed across academic leaders, faculty members, students, and quality professionals to strengthen quality literacy and increase the likelihood that evaluation findings inform pedagogical and curricular decisions. Third, universities should establish closed-loop improvement systems in which quality data are systematically translated into action plans, implemented through targeted interventions,

and subsequently re-evaluated using valid outcome measures. Fourth, greater attention should be given to the development of direct indicators of student learning outcomes, thereby reducing overreliance on perception-based measures and strengthening the evidence base for educational improvement.

Taken together, the review suggests that the effectiveness of IQA depends less on the presence of formal quality assurance structures than on the quality of their implementation and integration into organisational learning processes. Institutions that align improvement-oriented purposes, stakeholder engagement, and actionable quality processes are more likely to realise meaningful educational benefits than those relying primarily on compliance-driven quality assurance practices. Consequently, the future of IQA in higher education may lie not in expanding quality procedures, but in strengthening the mechanisms through which quality evidence is converted into sustained improvements in teaching and learning. Based on the synthesis, a practical design principle for effective IQA is the alignment of improvement-oriented purposes, distributed stakeholder ownership, and evidence-to-action feedback loops.

REFERENCES

- Afiralda, I., Bakri, M., Dewi, M. S., Nur, S., & Radzuanuddin, A. (2025). *Internal Quality Assurance System in Indonesian Basic Education : A Systematic Literature Review*. 8(2), 282–296. <https://doi.org/https://doi.org/10.31764/ijeca.v8i2.32857>
- Benson, H., Williams, K., & Heggart, K. (2024). *Quality assurance interventions in blended learning design : A systematic review of the literature*. 40(5), 31–46. <https://doi.org/https://doi.org/10.14742/ajet.9362>
- Erittu, E., & Turri, M. (2023). Internal quality assurance units: empirical evidence from Italy. *Quality in Higher Education*, 29(3), 358–373. <https://doi.org/10.1080/13538322.2022.2121197>
- Freeth, D., & Reeves, S. (2004). Learning to work together: using the presage, process, product (3P) model to highlight decisions and possibilities. *Journal of Interprofessional Care*, 18(1), 43–56. <https://doi.org/DOI: 10.1080/13561820310001608221>
- Gray, D. M., Bennett, M., Smart, K., & Kumar, A. (2022). An analysis of university academic department bylaws to measure commitment to program assessment: A U.S. case study. *Cogent Education*, 9(1). <https://doi.org/10.1080/2331186X.2022.2113218>
- Hoare, A., Dishke Hondzel, C., Wagner, S., & Church, S. (2024). A course-based approach to conducting program review. *Discover Education*, 3(1). <https://doi.org/10.1007/s44217-023-00085-4>
- Hou, A. Y. C., Chen, Y., Lin, A. F. Y., Su, E. H. C., Zhou, K. Z.-W., & Tao, C. H.-Y. (2025). Did Global Rankings and National Accreditation Drive Taiwanese Universities to Advance Sustainable Development Goals (SDGs)?—Competitive Advantage, a Moral Symbol or Leading to a Paradigm Shift. *Higher Education Quarterly*, 79(3). <https://doi.org/10.1111/hequ.70031>
- Iqbal, T., Ahmad, S., Aftab, F., & Mahmood, C. K. (2025). Enhancing higher education institutions' performance: The mediating role of academic accreditation in quality management initiatives in UAE. *SAGE Open*, 15(3), 21582440251358980. <https://doi.org/https://doi.org/10.1177/2158244025135898>
- Knyaginina, N. V., & Puchkov, E. V. (2024). Internal Education Quality Assurance Systems in the Documents of Russian Universities: Quality Assurance or its Appearance? *Voprosy Obrazovaniya/Educational Studies Moscow*, 2(3). <https://doi.org/https://doi.org/10.17323/vo-2024-17221>
- Krooi, M., Whittingham, J., & Beausaert, S. (2024). *Studies in Educational Evaluation Introducing the 3P conceptual model of internal quality assurance in higher education : A systematic literature review*. 82(April), 1–10. <https://doi.org/DOI: 10.1016/j.stueduc.2024.101360>

- Leiber, T., & Seyfried, M. (2025). The role of quality literacy in understanding quality in higher education. *Quality in Higher Education*, 31(2), 268–283. <https://doi.org/DOI:10.1080/13538322.2025.2532981>
- Lou, L. (2025). From Fragmentation to Coherence: A Knowledge Graph-Driven Integration Framework for Blended Learning. *Arab World English Journal (AWEJ) Volume*, 16(3), 211–226. <https://doi.org/doi:https://dx.doi.org/10.24093/awej/vol16no3.12>
- Mazza, T., & Azzali, S. (2025). Stakeholders' involvement in economics and management programs quality assurance. *Quality Assurance in Education*, 33(1), 47–61. <https://doi.org/DOI:10.1108/QAE-11-2023-0193>
- Miguel, L., Martins, C., & Domingues, J. P. (2019). *ISO 9001 : 2015 Adoption : A Multi-Country Empirical Research*. 12(1), 27–50. <https://doi.org/https://doi.org/10.3926/jiem.2745>
- Mukhatayev, A., Omirbayev, S., Kassenov, K., & Idiyatova, Y. (2024). Quality assurance system of higher education in kazakhstan through stakeholders' eyes: An empirical study to identify its challenges. *Education Sciences*, 14(12), 1–36. <https://doi.org/https://doi.org/10.3390/educsci14121297>
- Pham, N. T. T., Nguyen, C. H., Pham, H. T., & Ta, H. T. T. (2022). Internal quality assurance of academic programs: A case study in Vietnamese higher education. *Sage Open*, 12(4), 1–11. <https://doi.org/https://doi.org/10.1177/21582440221144419>
- Pramono, S. E., & Widiyanto, W. (2024). Quality assurance in higher education: Trends and insights from a systematic literature review. *Edelweiss Applied Science and Technology*, 8(6), 6579–6588. <https://doi.org/https://doi.org/10.55214/25768484.v8i6.3410>
- Pujianti, Y., Aminah, S., Nuryati, E., Mulyanto, A., & Sariff, A. F. (2025). A Systematic Literature Review of Internal Quality Assurance in Early Childhood Education in Developed and Developing Countries: Internal Quality Assurance in Early Childhood Education. *Indonesian Journal of Educational Research and Review*, 8(1), 273–291. <https://doi.org/https://doi.org/10.23887/ijerr.v8i1.90137>
- Rahman, M. M., & Nasrin, S. (2024). Perceived service quality at higher education institutions: A study on the success factors of total quality management practices in Bangladesh. *Social Sciences and Humanities Open*, 10. <https://doi.org/10.1016/j.ssaho.2024.100997>
- Serrano, O. M. E., Miranda González, F. J., & Mourato, J. A. B. (2025). Determinants of the effectiveness of quality assurance systems and institutional performance in higher education. *Cogent Education*, 12(1). <https://doi.org/10.1080/2331186X.2025.2581411>
- Širec, K., & Rožman, M. (2025). Collaborative Leadership for Quality Assurance: A Case Study on Developing a Strategic Quality Manual in Higher Education. *Education Sciences*, 15(12), 1627. <https://doi.org/https://doi.org/10.3390/educsci15121627>
- Sofyani, H., Abu Hasan, H., & Saleh, Z. (2023). Does internal control contribute to quality management in higher education institutions? Indonesia's adoption experience of the COSO integrated framework. *The TQM Journal*, 35(8), 2162–2180. <https://doi.org/https://doi.org/10.1108/TQM-06-2022-0201>
- Stalmeijer, R. E., Whittingham, J. R. D., Bendermacher, G. W. G., Wolfhagen, I. H. A. P., Dolmans, D. H. J. M., & Sehlbach, C. (2023). Continuous enhancement of educational quality–fostering a quality culture: AMEE Guide No. 147. *Medical Teacher*, 45(1), 6–16. <https://doi.org/https://doi.org/10.1080/0142159X.2022.2057285>
- Vargas, H., Heradio, R., Farias, G., Lei, Z., & de la Torre, L. (2024). A Pragmatic Framework for Assessing Learning Outcomes in Competency-Based Courses. *IEEE Transactions on Education*, 67(2), 224–233. <https://doi.org/10.1109/TE.2023.3347273>
- Winstone, N. E., & Carless, D. (2021). Who is feedback for? The influence of accountability and quality assurance agendas on the enactment of feedback processes. *Assessment in Education: Principles, Policy and Practice*, 28(3), 261–278. <https://doi.org/10.1080/0969594X.2021.1926221>
- Wysocka, K. (2022). Internationalization and Quality Assurance in Higher Education. *Internationalization and Quality Assurance in Higher Education*, 26(1), 204–230. <https://doi.org/10.2478/manment-2019-0091>

Zheng, X., Vastrad, S., He, J., & Ni, C. (2023). Contextualizing gender disparities in online teaching evaluations for professors. *PLoS One*, 18(3), 1–24.
<https://doi.org/https://doi.org/10.1371/journal.pone.0282704>