International Journal of Evaluation and Research in Education (IJERE)
Vol. 15, No. 4, August 2026, pp. 2840~2851
ISSN: 2252-8822, DOI: 10.11591/ijere.v15i4.40024

Reframing teacher evaluation in higher education: a three-pillar framework from Assam

Arabinda Rajkhowa, Munmi Dutta
Department of Assamese, North Lakhimpur University, North Lakhimpur, India

Article Info

Article history:

Received Apr 17, 2026
Revised Jun 26, 2026
Accepted Jul 3, 2026

Keywords:

IQAC
Peer review
Quality assurance
Self-evaluation
Teacher evaluation

ABSTRACT

Teacher evaluation shapes the quality of classroom instruction and, through it, student learning outcomes; yet in Indian higher education the dominant single-source model, student feedback channeled through the internal quality assurance cell (IQAC), is widely critiqued as ritualistic and developmentally inert. To the authors' knowledge, no prior study has integrated student feedback, structured self-evaluation, and peer review into a coherent operational framework for regional Global South contexts. Drawing on primary data collected between 2009 and 2025 from approximately 200 undergraduate arts and science students, principally at North Lakhimpur College (now North Lakhimpur University) and other institutions across Assam, this study employs a qualitative-descriptive design with thematic analysis and triangulation of open-ended questionnaires and semi-structured interviews. Two contributions emerge. First, it yields a culturally grounded fivefold taxonomy of good teaching from Assamese student articulations, including culturally distinctive expectations: the teacher's public moral role in the community and the obligation of intellectual life beyond the syllabus, that standardized student evaluation of teaching (SET) instruments routinely miss. Second, it proposes a developmental three-pillar framework integrating reformed student feedback, disciplined teacher self-evaluation, and structured peer review to restore the formative function of evaluation and improve student learning outcomes. Both the taxonomy and the framework are scalable across comparable institutions in the Global South.

Corresponding Author:

Arabinda Rajkhowa
Department of Assamese, North Lakhimpur University
North Lakhimpur, Assam 787031, India
Email: arabinda@nlu.ac.in

1. INTRODUCTION

The question of how to evaluate teaching in higher education has gained renewed salience in the 21st century. Rising enrollment, diversification of institutional forms, and the integration of national education systems into global quality-assurance frameworks have put sustained pressure on universities and colleges to demonstrate, measure, and improve the quality of their teaching [1]. The validity of student evaluation of teaching (SET) as a summative measure has been extensively contested [2]. Meta-analytic evidence challenges the assumption that more highly rated instructors produce greater student learning [3], [4], and a substantial body of work documents systematic biases against women and instructors from minority backgrounds [5]–[11]; on the strength of this evidence, some scholars now argue that SET should be abandoned as a summative measure altogether [12].

In India, this pressure has taken institutional form through the internal quality assurance cell (IQAC) at every college accredited by the National Assessment and Accreditation Council (NAAC) [13], which

operates as an autonomous body under the University Grants Commission (UGC). The policy environment has been substantially reshaped by the National Education Policy (NEP) 2020 [14], which explicitly identifies continuous professional development, peer review, and outcome-based evaluation of teachers as central pillars of higher-education reform. The UGC regulations [15] governing teacher appointments and promotions specify performance-based criteria that presuppose a functioning institutional evaluation mechanism.

Despite this elaborate architecture, the operational reality in many affiliated colleges, particularly in the less urbanized districts of northeastern India, remains far from the intended design. Student-based appraisal of teaching, the most common IQAC instrument, often degenerates into a routine paper exercise: forms are distributed openly, collected without adequate safeguards for anonymity, filed away, and rarely followed up with the teacher. The consequences are not procedural but substantive: many teachers complete entire careers without ever receiving credible developmental feedback on their teaching; early-career instructors have no mechanism to correct ineffective practice; pedagogical stagnation persists unchallenged; and the aggregate quality of classroom instruction, and therefore student learning, is impaired. This directly affects student engagement, comprehension, and long-term learning outcomes. This operational gap is the context for the present study.

Although student ratings are imperfect, student perceptions remain an important source of information about teaching quality. Alhija [16] identifies a consistent cluster of attributes: clear explanation, organization, respect, availability, and knowledge, that students across contexts associate with good teaching. Catano and Harvey [17] developed and validated the evaluation of teaching competencies scale, yielding seven dimensions with strong reliability. Mimirinis and Ahlberg [18] used phenomenographic analysis to examine doctoral students' conceptions of university teaching, demonstrating a hierarchical progression from transmission-oriented to transformation-oriented understandings.

The literature on teacher self-evaluation and reflective practice has grown substantially in recent years. Blumberg [19] shows that critical reflection helps faculty implement learner-centered teaching by revealing the gap between espoused and enacted pedagogy. Myllykoski-Laine *et al.* [20] report on the use of the research-based HowUTeach self-reflection tool in Finnish higher education and find that structured prompts, coupled with feedback, promote deeper reflection-on-action. Constantinou and Wijnen-Meijer [21], focusing on medical schools, argue for triangulating SET data with peer review, program administrator input, and self-evaluation, providing a multi-source template that has not yet been operationalized in regional Indian higher education.

Three gaps in the existing literature are evident. First, although critiques of SET, studies of self-evaluation, and proposals for peer review exist, no study has integrated all three into a coherent, operational evaluation framework calibrated to the resource-constrained, semi-urban higher-education contexts that characterize much of regional India and the broader Global South. Second, the critique of student-only evaluation models has been developed primarily from European and North American data; its applicability to Indian institutional realities, where internal-assessment marks are awarded by the same teacher being evaluated, has not been examined. Third, although NEP 2020 mandates peer review and continuous professional development, no district-level empirical research has translated these national directives into implementable routines for IQAC coordinators, even though national policy without local operational evidence remains hollow.

This study focuses on the higher-education landscape of Assam, with North Lakhimpur College, a well-established undergraduate institution in Lakhimpur district that the Government of Assam upgraded to university status as North Lakhimpur University in 2023, as the principal site of investigation, complemented by data gathered from other higher educational institutions in different parts of the state. Initiated at the principal site in 2009 and continued through 2025, the study draws on evidence accumulated over more than a decade and a half. Lakhimpur shares with much of rural and semi-urban Assam a set of features that make it an analytically rich and replicable case: undergraduate colleges drawing students largely from first-generation learner backgrounds, a teaching workforce historically underpaid and often working for extended periods without provincialized salaries before government absorption, and a public culture in which the social authority of the teacher has been visibly eroding. These features are common across northeastern India and recur in comparable institutions elsewhere in the Global South, so evidence and remedies generated here are scalable beyond the immediate site.

Against this background, the study is guided by three research questions that progress from operational diagnosis to constructive synthesis. The first question identifies the operational deficiencies in the current evaluation regime as experienced by students at the research site. The second question turns to the constitutive expectations students bring to good teaching, while the third asks how multi-source evaluation can be institutionalized to overcome the limitations of single-source models. The three questions are stated in their final form:

- What operational deficiencies exist in IQAC-mediated student evaluation as currently practiced in undergraduate colleges in Assam? (RQ1)
- What attributes do students in this context identify as constitutive of a good college teacher? (RQ2)

- How can teacher self-evaluation, integrated with reformed student feedback and peer review, be institutionalized to address the limitations of single-source evaluation? (RQ3)

The novelty of this study lies in integrating three elements not previously combined in research on regional Indian higher education. First, it presents one of the earliest systematic empirical accounts of IQAC-mediated teacher evaluation in undergraduate colleges in Assam, addressing a regional context that has been largely absent from the international literature on teacher evaluation. Second, and equally important, it advances a fivefold, student-derived taxonomy of effective teaching grounded in the Assamese cultural setting rather than imposed through pre-structured Western rating instruments; this taxonomy is a substantive empirical contribution in its own right, since it incorporates culturally distinctive expectations, most notably the teacher's public moral role in the surrounding community and the expectation of intellectual life beyond the syllabus, that are routinely missed by standardized SET instruments. Third, the study proposes a developmental three-pillar framework that integrates reformed student feedback, disciplined teacher self-evaluation, and structured peer review into a single operational design: a rare, integrated operational design calibrated for resource-constrained, semi-urban higher-education contexts in the Global South and explicitly translating NEP 2020 directives into implementable district-level practice. Taken together, these contributions move the discussion of teacher evaluation in Indian higher education from generalized critique to a tested, replicable framework supported by a culturally grounded empirical taxonomy and concrete operational specifications.

2. METHOD

2.1. Research design

The study adopts a qualitative-descriptive design with a primarily analytical orientation. A qualitative approach was considered appropriate because the study sought to capture the nuanced, lived experiences of students with their teachers, rather than generate statistical generalizations about a large population. The descriptive component was retained to enable orderly categorization of student responses into analytically useful themes. This design aligns with recent methodological guidance in the field, which increasingly favors qualitative and mixed-methods approaches to teaching-quality research to overcome the well-documented limitations of purely quantitative rating scales [21].

A Likert-scale instrument was deliberately rejected for two reasons. First, pre-structured Likert items would have replicated the very deficiencies this study critiques: compressed response variance, suppression of culturally specific concepts such as the teacher's social role, and the conflation of likability with pedagogical competence [22]. Second, the study sought authentic articulations of student experience, which open-ended formats are uniquely positioned to elicit. Therefore, the qualitative-descriptive approach is the most suitable method for the research problem at hand.

2.2. Site and participants

The principal research site was North Lakhimpur College, a prominent undergraduate institution in Lakhimpur district, Assam, which the Government of Assam upgraded to university status as North Lakhimpur University in 2023. The study was initiated at this institution in 2009 and continued through 2025; over this period, data were also collected from a number of other higher educational institutions located in different parts of Assam, so that the evidence base extends beyond a single campus. The principal site was chosen for three reasons: it is representative of district-level higher education across northeastern India in its student demographics and faculty workforce profile; both authors hold teaching and research positions at the institution and could therefore secure informed institutional consent and classroom access; and the institution has a fully implemented IQAC, making it an appropriate site for examining IQAC-mediated evaluation in practice. Taken together, the principal site and the supplementary institutions establish an analytically rich and methodologically appropriate evidence base whose findings are reasonably transferable to comparable institutions across northeastern India and the broader Global South.

Students in the arts and science streams enrolled in the three-year degree program, across first-, second-, and third-year cohorts, formed the participant pool, drawn principally from the main site and supplemented by respondents from the other participating institutions in Assam. Sampling was purposive rather than random: heads of departments facilitated access to classrooms during designated hours, and participation was voluntary. Approximately two hundred students were approached for written responses, and from these, a subset of about one hundred and eighty fully completed submissions was retained for thematic analysis. The sample size was chosen to allow saturation of themes within each stream while remaining manageable for close qualitative reading; the retained number comfortably exceeds the saturation threshold typically reported for thematic studies of pedagogical perception. All respondents were informed in writing that their identities would not be recorded and that submissions would be used solely for research purposes.

2.3. Instruments

Two instruments were used in combination. The first was an open-ended written questionnaire inviting students to describe, in their own words: i) the qualities essential to a good teacher; ii) the limitations they had experienced in their teachers; iii) their experience with the formal teacher evaluation forms administered through the IQAC; and iv) their suggestions for improving the teaching they received. The second was a set of semi-structured interviews conducted with a smaller subset of students who had submitted particularly thoughtful written responses, used to probe specific themes in greater depth. The combination of written and oral instruments enabled methodological triangulation of findings within the qualitative paradigm.

2.4. Data collection procedure

Written responses were collected in classroom settings during regular college hours, with prior institutional permission, over an extended period spanning 2009 to 2025 at the principal site and the additional participating institutions. Students were given approximately ninety minutes to submit their responses. No teacher of any respondent was present during the collection to ensure that the power asymmetry between teacher and student would not influence what students wrote. Responses were sealed in envelopes and returned directly to the researcher. Interviews were conducted in a separate session, with audio-recorded consent, and were subsequently transcribed for analysis. All data were anonymized at the point of collection.

2.5. Analytical approach

The analysis followed a thematic-analytic logic. The researchers first read the entire set of responses to develop a broad sense of the data. In a second reading, recurring statements about teacher qualities, teacher deficiencies, and students' experiences with formal evaluation were coded. These codes were then grouped into higher-order themes through successive iterations. Five dominant themes emerged regarding the attributes of a good teacher, and five regarding the deficiencies of IQAC-mediated evaluation. Interview material was used primarily to deepen, qualify, and occasionally nuance the written data.

2.6. Trustworthiness and credibility

Several procedures were implemented to strengthen the credibility and trustworthiness of the findings. First, both authors independently coded a randomly selected subset of responses and then compared codes; disagreements were resolved through discussion, and the coding scheme was iteratively refined until inter-coder agreement was satisfactory. Second, written and interview data were triangulated to confirm or qualify emerging themes. Third, member-checking was conducted during the interview phase: students were asked whether the preliminary themes derived from the written responses captured their experience, and their feedback was incorporated. Fourth, an audit trail of code-to-theme decisions was maintained throughout the analysis. Fifth, a reflexivity statement was developed by the authors: as faculty members at the research site, the authors recognized the risk of insider bias and explicitly bracketed their evaluative judgments during coding. Anonymity was preserved at every stage, and the data set contains no personal identifiers. These procedures ensure methodological rigor and trustworthiness consistent with qualitative research standards.

3. RESULTS

This section presents the study's empirical findings without interpretation. This section addresses RQ1 by characterizing the operational deficiencies of IQAC-mediated evaluation as reported by students and RQ2 by outlining the qualities of a good teacher articulated by the same respondents. The corresponding interpretation, comparison with the literature, and the integrated framework that addresses RQ3 are presented separately.

3.1. Operational deficiencies of IQAC-mediated evaluation (RQ1)

Five consistent patterns emerged from student responses to the IQAC-based evaluation they had encountered. Each pattern is reported as a pure empirical finding, with supporting evidence drawn directly from the written and interview data. The interpretation of these findings in dialogue with the existing literature is presented separately. First, a large proportion of respondents described the evaluation exercise as a mere formality. Students reported that forms were distributed without adequate explanation, completed quickly, and collected en masse. When asked what happened to their responses, very few could identify any follow-up: there was no report, no perceptible change in teacher behavior, and no acknowledgment of the themes raised.

Second, the lack of confidentiality was widely perceived. Forms were often distributed and collected by a staff member or student volunteer known to respondents. Even when names were not recorded, students

reported that handwriting, stream, section, and seat order made identification possible in practice. In a district college with small classes and high teacher–student familiarity, nominal anonymity does not translate into perceived anonymity.

Third, internal assessment strongly influences student responses. Under the semester system in operation across Assam's degree colleges, between 20 and 30% of final marks are awarded by the subject teacher as internal assessment. Students are therefore rationally cautious about offering any assessment that could be traced back to them. As a result, reported ratings cluster at the top of the available scale and cease to serve any diagnostic function.

Fourth, students acknowledged that the judgment needed to evaluate a teacher is not equally developed in all of them. Several respondents noted that the "easy" teacher, one who jokes, relaxes the rules, or is lenient with grades, is immediately popular and likely to receive favorable ratings, whereas the demanding teacher who insists on standards may be rated less favorably until the student matures academically. Respondents indicated that this discrepancy in evaluative maturity often produces ratings that reward affability rather than pedagogical effort.

Fifth, the lack of review was conspicuous. Students reported that they had never been informed of aggregate findings, of any changes in teaching practice attributable to feedback, or of any mechanism for institutional learning from the exercise. In the absence of such returns, students reported they had no observable evidence that the evaluation had any consequence for the quality of teaching they received.

3.2. Student-articulated qualities of a good teacher (RQ2)

When students were asked to describe the qualities of a good teacher, five themes emerged consistently across both written and interview data. Table 1 summarizes these themes along with illustrative paraphrases drawn from student responses, and is followed by short empirical descriptions of each theme in turn. The interpretation of these themes in relation to existing literature is presented separately.

Table 1. Student-articulated qualities of a good college teacher (Lakhimpur, Assam)

Theme	Description	Representative student response (paraphrase)
Discipline and punctuality	Teacher models the regulative conduct expected of students: arriving on time, staying for the full class, and meeting deadlines.	*Our teachers reprimand us for arriving late, yet they enter ten minutes after the bell and leave five minutes early.*
Support outside the classroom	The teacher remains approachable beyond formal class hours—in corridors, offices, and during informal interactions.	*Inside class, he is helpful, but outside, he barely acknowledges us. We want a teacher who is the same in both places.*
Social role and conduct	The teacher accepts public responsibility in the surrounding community and speaks with dignity on matters of importance.	*When something important happens in our area, most teachers stay silent. We expect them to speak.*
Inspiration and praise	The teacher recognizes genuine student effort with timely, measured encouragement, avoiding performative praise.	*A single sincere word from a teacher can keep us working for weeks. Empty praise feels worse than silence.*
Knowledge beyond the syllabus	The teacher brings wider reading, reflection, and lived experience into the classroom rather than reciting textbook material.	*We do not come to class only for the textbook. When a teacher brings in something from a book or life, the class feels alive.*

Discipline and punctuality were the most frequently cited attributes. Respondents uniformly objected to teachers arriving late to class and leaving early, especially when those teachers enforced strict discipline on students for similar lapses. From the student perspective, punctuality is a basic contractual expectation that precedes all pedagogical considerations, and any visible breach of it was reported as undermining the teacher's authority before instruction had even begun.

Support outside the classroom emerged as a second theme. Students valued teachers who remained approachable, willing to answer questions, and available in corridors, offices, or other settings. Teachers who abandoned cordiality outside the classroom were perceived as fragmented, and several respondents explicitly contrasted classroom warmth with corridor coldness as a sign of inauthentic pedagogical care.

Social role and conduct were a distinctive and recurring concern. Students in Lakhimpur articulated a clear expectation that a teacher should occupy a visible, responsible position in the surrounding community: the teacher's voice should be heard on matters of public importance. When teachers remained silent or detached, respondents expressed disappointment and reported that such silence diminished the teacher's standing in their eyes regardless of classroom performance.

Inspiration and praise formed a fourth theme. Students valued recognition of good work: a single word of encouragement, acknowledgment of sustained effort, or a recommendation for further study. At the same time, they were sensitive to the difference between genuine praise and performative praise offered to

demonstrate the teacher's magnanimity, and they reported a clear preference for measured, specific acknowledgment over generalized flattery.

Knowledge beyond the syllabus was the fifth, and perhaps the most intellectually demanding, expectation. Students reported that they did not expect merely the textbook; rather, they expected their teacher to bring into the classroom material drawn from reading, reflection, and life experience to enrich the subject. Teachers who limited themselves to reciting notes were quickly seen as limited, and their authority eroded in the respondents' perception.

4. DISCUSSION

This section interprets the empirical findings in dialogue with the existing literature, develops the enabling conditions for credible self-evaluation, and presents the integrated framework that addresses RQ3. The study therefore does not reject student perception as a source of evaluative insight; rather, it argues that student feedback becomes developmentally credible only when triangulated with disciplined teacher self-evaluation and structured peer review. The subsections interpret the deficiencies and the student-articulated qualities, respectively; the section then articulates the framework its operational specification, concrete recommendations for IQAC coordinators, and its practical implications.

4.1. Interpreting the operational deficiencies in relation to existing literature

The five operational deficiencies can be interpreted as a coherent pattern when read against the international literature on SET, and the resulting interpretation directly informs the design of the integrated framework. Before the individual findings are taken up, it should be stressed that the study does not reject student perception as a source of evaluative insight, but argues that student feedback becomes developmentally credible only when triangulated with disciplined teacher self-evaluation and structured peer review. Each finding is taken up in turn below, with explicit attention to whether it corroborates, extends, or differs from prior work, and to the gap it helps address.

The finding that evaluation is widely experienced as a mere formality corroborates Constantinou and Wijnen-Meijer [21], who note that when students do not see their feedback acted upon, their motivation to provide sincere responses drops, producing a self-reinforcing cycle of instrument decay. It also deepens Zhao *et al.* [23], who describe ritualization in Chinese and Western higher education, by offering district-level confirmation that the same dynamic operates structurally in Indian undergraduate colleges. The implication for the present study is that the IQAC must be reconfigured to function as an analytical node that returns findings to teachers, rather than as a passive collector of forms, and this implication directly motivates Pillar 1 of the integrated model.

The finding on confidentiality adds district-level evidence to Quansah *et al.* [1], who identify confidentiality as a necessary condition for candid response, and reveals that Indian implementation has not operationalized this condition in any substantive sense. It contributes to the gap identified in the literature review by illustrating that the well-documented confidentiality requirement assumes anonymity-in-practice, not merely anonymity-in-form, and that small-class district colleges fall structurally short on this requirement. The finding therefore reinforces the need for online or sealed-box collection mechanisms detailed in the operational specification of Pillar 1.

The finding on internal-assessment influence further demonstrates the dynamics analyzed by several researchers [24], [25]: in semester-system colleges where internal-assessment marks are awarded by the same teacher being evaluated, the threat-response dynamic operates structurally rather than through individual instructor incentives. This is an explicit contribution to the research gap identified earlier, since prior critiques of evaluator–evaluatee role overlap have been developed primarily on European and North American data, and the present study reveals how the dynamic intensifies in the Indian semester system. The implication is that the timing of feedback collection must be deliberately repositioned.

The finding that student ratings reward affability corroborates previous studies [3], [22], whose combined argument is that SET often measures likability rather than pedagogically relevant competence. The present study adds the methodologically valuable observation that students themselves recognize and articulate this tendency, which illustrates that single-source student evaluation cannot be relied upon to supply pedagogically diagnostic information even when sincerity is assumed. This finding helps justify the integration of peer review and self-evaluation alongside reformed student feedback, as elaborated in the integrated model.

The finding on the absence of feedback loops adds empirical weight to the distinction that Spooren *et al.* [2] draw between formative and summative uses of SET. Their argument that the formative use requires feedback loops finds direct empirical confirmation in the present data, in which no respondent could identify any institutional learning from the exercise. This finding contributes to the research gap by

revealing that the formative function of evaluation, often assumed in policy documents, is structurally absent at the district level, and that restoring it requires institutional redesign rather than only better instruments.

4.2. Interpreting the student-articulated qualities of a good teacher in relation to existing literature

The five student-articulated qualities are interpreted here in dialogue with the international literature on good teaching, with each theme considered separately. Where the present findings corroborate prior work, they are noted as confirmatory; where they deepen or differ from it, they are flagged accordingly, with explicit attention to their contribution to the research gap. The interpretive paragraphs that follow are organized in the same sequence, so that the parallels between empirical findings and their interpretation are easy to trace.

The discipline-and-punctuality theme aligns with Alhija [16] and the validated factor of classroom organization in Catano and Harvey [17]. The contribution of the present study lies in revealing that students treat punctuality as a contractual precondition rather than as one ratable competence among others, which has direct implications for how organization is captured in any reformed student-feedback instrument. The theme therefore corroborates prior work while adding this contractual framing.

The support-outside-the-classroom theme aligns with Witter and Hattie [26], who report that teachers scoring high on self-efficacy and intrinsic motivation extend their cooperative disposition beyond the classroom and that students perceive this as an integrated personal quality rather than a professional gesture. The Lakhimpur data deepen this finding by revealing that students penalize incongruence between inside-class and outside-class conduct in their evaluative judgments. This contributes to the research gap by indicating that any credible peer-observation rubric must include availability beyond the timetabled session as a dimension of teaching quality.

The social-role-and-conduct theme is culturally inflected and aligns with Kulal *et al.*'s [27] account of the Indian teacher's traditional standing as a moral exemplar; it also deepens Kishimoto [28] on faculty self-reflection beyond the classroom into civic life. The present study's contribution is to ground this expectation in primary district-level data, which has been substantially absent from prior literature on Indian higher education. This theme differs from the cluster of attributes reported by Alhija [16] in Western contexts and therefore documents a culturally specific dimension of good teaching that pre-structured Western instruments would have missed.

The inspiration-and-praise theme aligns with findings in medical education [21] and with the phenomenographic hierarchy of Mimirinis and Ahlberg [18]. The fine-grained discrimination that respondents drew between genuine and performative praise corroborates the higher levels of that hierarchy and reveals that students operate with a more sophisticated evaluative vocabulary than coarse Likert items can capture. The contribution to the research gap is the demonstration that this discrimination is articulable in open-ended formats and should therefore inform the design of qualitative-feedback prompts in reformed student-evaluation instruments.

The knowledge-beyond-the-syllabus theme corroborates Mimirinis and Ahlberg [18] and the broader reflective-practice argument [19], [29] that the reading teacher is the precondition for the thinking classroom. The Lakhimpur data add contextual evidence to this work by documenting that students recognize and explicitly value such intellectual breadth even in resource-constrained district institutions where it might be expected to be diluted. This contribution to the research gap reinforces the design of the self-evaluation prompts in Pillar 2, particularly prompt 8 on recent disciplinary reading.

4.3. Enabling conditions for credible self-evaluation

The analysis suggests that credible institutional teacher self-evaluation requires several enabling conditions. First, the teacher must have the capacity for dispassionate self-analysis: the ability to move beyond defensive reactions and view shortcomings as opportunities for development. Second, the teacher must have access to evidence of former students' accomplishments and trajectories, which serve as delayed, longitudinal indicators of teaching impact unavailable through end-of-course ratings. Third, the teacher must hold an articulable ideal of teaching, a conception of what teaching is for, against which practice can be measured. Fourth, the teacher must adopt a professional stance vis-à-vis ethnic and religious identity, resisting the temptation to import particularistic identities into the classroom. Finally, the teacher must engage in sustained reading. The literature on reflective practice [19], [20], [29]–[32] supports each of these conditions and provides instruments, such as HowUTeach, that can operationalize them for routine use.

4.4. The three-pillar framework: integrating student feedback, self-evaluation, and peer review (RQ3)

By synthesizing empirical findings with the literature, the analysis points to a developmental three-pillar framework for teacher evaluation, as illustrated in Figure 1. The framework is explicitly growth-oriented rather than compliance-oriented: its purpose is to generate actionable feedback that improves

classroom practice and, ultimately, student learning outcomes. The first pillar is credible, confidential student feedback with genuine institutional follow-up: an instrument reformed to strengthen confidentiality protocols, timed to minimize internal-assessment distortion, and linked to an analytical routine that returns findings to the teacher. The second pillar is disciplined, periodic teacher self-evaluation grounded in reading, observation of former students, and articulated teaching ideals. The third pillar is structured peer review, including classroom observation by senior colleagues, which supplies the pedagogical expertise that students necessarily lack. Constantinou and Wijnen-Meijer [21] make this argument essentially for medical education, and its logic transfers to undergraduate colleges. Within such an arrangement, the IQAC functions not as a collector of forms but as the analytical node that integrates three data streams and returns structured, developmentally oriented feedback. As shown in Figure 1, the integration of the three pillars feeds directly into improved teaching quality and, in turn, enhanced student learning outcomes. Such a reconfiguration is consistent with the directions of NEP 2020 [14] and the quality-enhancement mandate of NAAC [13]. The integrated model therefore functions as the central organizing structure of the study's contribution.

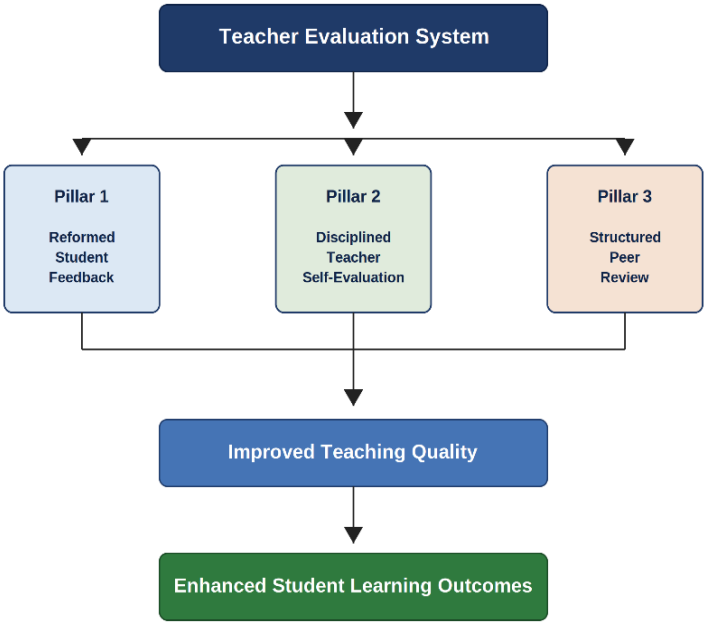

Figure 1. Three-pillar framework for teacher evaluation

4.5. Operational specification of the three pillars

4.5.1. Pillar 1: reformed student feedback

Three operational reforms make student feedback credible. First, anonymity must be protected through online or sealed-box submission, with no intermediary known to the respondent, so that the respondent perceives anonymity rather than merely receiving an assurance of it. Second, the evaluation cycle must be timed after internal-assessment marks are finalized to break the strategic-response loop. Third, aggregated findings must be returned to the teacher with structured analytical commentary within a fixed institutional timeline, typically four weeks after collection, so the loop closes and the instrument's formative function is restored.

4.5.2. Pillar 2: disciplined self-evaluation

To make self-evaluation operational rather than aspirational, every teacher should complete a structured annual self-evaluation using concrete reflective prompts. The prompts are deliberately framed to invite candid acknowledgment of weaknesses as well as strengths, in line with the reflective-practice literature. The following eight prompts, drawn from the present analysis and that literature, are recommended for institutional adoption:

- In which classes this semester did I feel my explanation was inadequate, and what was missing?
- Did I engage students effectively, and what evidence (questions, attention, and follow-up) supports my judgment?
- Did I go beyond the syllabus, and on which topics am I still teaching from notes I prepared during my own student years?

- Whose sustained academic trajectory among my former students reflects something I taught well, and what was it?
- Where did my classroom behavior diverge from the conduct I expect from students (punctuality, preparation, and civility)?
- Did I import any non-pedagogical identity (regional, political, and religious) into the classroom this term, and with what effect on students?
- Which student feedback theme from this year's evaluation surprised me, and how do I plan to address it?
- Which book or article in my discipline have I read in the past six months that has changed how I teach?

These prompts may be completed in writing and retained by the teacher for their own developmental use. The only mandatory institutional requirement should be the submission of the completed form, not the disclosure of its content, in order to protect the candor of self-reflection. This arrangement preserves the formative purpose of self-evaluation while still providing the IQAC with verifiable evidence of compliance for accreditation purposes.

4.5.3. Pillar 3: structured peer review

To prevent peer review from becoming either a clique-based ritual or a punitive instrument, the following operational safeguards are recommended. Observation pairs should be assigned across departments rather than within them to limit in-group effects. Each peer review cycle should include two scheduled observations by the reviewer, separated by at least four weeks, to capture variation in classroom practice. Observations should follow a pre-agreed rubric covering four dimensions: clarity of explanation, classroom organization, student engagement, and alignment of session content with stated learning outcomes. Reviewer training, comprising at least one half-day workshop, should be a precondition for serving as a peer observer. Feedback should be developmental in form ("two strengths, one growth area") rather than evaluative for promotion decisions, and should be returned in writing to the observed teacher within seven days, with the observed teacher's right to a written response on the record. These safeguards collectively address the risk of bias and arbitrariness that have historically prevented peer review from gaining institutional acceptance in Indian colleges.

4.6. Recommendations for IQAC coordinators

Three concrete, actionable recommendations follow directly from the preceding analysis and discussion and are offered to IQAC coordinators at undergraduate colleges for adoption. Each recommendation addresses one of the operational deficiencies, while at the same time advancing one pillar of the integrated model set out. They are presented in the order in which they would typically be implemented in an annual institutional cycle.

- Recommendation 1: reschedule student feedback collection. Move the IQAC student-feedback cycle to the period after internal-assessment marks have been submitted to the controller of examinations, thereby structurally minimizing strategic student responses. Publish the revised collection date in the academic calendar at the start of the year, and ensure that responses are aggregated and returned to teachers within four weeks.
- Recommendation 2: institute mandatory teacher self-evaluation logs. Adopt the eight-prompt self-evaluation form as an annual institutional requirement, and have the IQAC verify submission of the completed form (but not its content). The form should be retained by the teacher for personal development, with institutional access limited to summary aggregates that protect candor.
- Recommendation 3: implement structured peer-observation cycles. Establish annual cross-departmental peer-observation pairs under the operational safeguards, with a published rubric, mandatory observer training, and a fixed feedback timeline. Peer-observation outputs should feed into the IQAC's annual analytical report and inform institutional faculty development planning.

4.7. Practical implications

The findings have practical implications for three stakeholder categories. For teachers, the framework offers a developmental rather than punitive instrument: candid self-evaluation, supplemented by trustworthy student feedback and constructive peer review, provides the evidence base for sustained pedagogical growth. For institutions, the model transforms the IQAC from a document collector into an analytical node that delivers actionable insights to faculty within fixed timelines, increasing the developmental yield of existing accreditation-related infrastructure. For policymakers and accreditation agencies, the proposed evaluation model operationalizes NEP 2020's otherwise abstract directives on outcome-based evaluation and continuous professional development and provides a replicable template for regional implementation that does not depend on the resource intensity of metropolitan institutions. Improved

teaching quality at scale, the eventual consequence, is expected to translate into measurable improvements in student learning outcomes. Taken together, these implications further reinforce the case for the integrated model as a coherent, developmental alternative to the prevailing single-source evaluation approach.

5. CONCLUSION

This study examined IQAC-mediated teacher evaluation in undergraduate colleges in Assam, principally at North Lakhimpur College, now North Lakhimpur University, and produced two complementary contributions. The empirical analysis identified five operational deficiencies in single-source evaluation: ritualization, weak confidentiality, internal-assessment distortion, the rewarding of affability over pedagogical effort, and the absence of feedback loops. It also yielded a culturally grounded fivefold taxonomy of good teaching in which culturally distinctive expectations, namely the teacher's public moral role in the community and the obligation of intellectual life beyond the syllabus, feature alongside discipline, availability, and pedagogically meaningful praise. These findings were translated into a developmental three-pillar framework integrating reformed student feedback, disciplined teacher self-evaluation, and structured peer review, operationalized through eight self-evaluation prompts, three IQAC recommendations, and concrete peer-review safeguards.

The significance of these contributions is both empirical and practical. The fivefold taxonomy documents culturally distinctive dimensions of good teaching that pre-structured Western instruments routinely miss and that have not previously been captured at the district level in Indian higher education, while the three-pillar framework moves evaluation discourse from generalized critique to a tested, replicable design that translates the abstract directives of NEP 2020 into implementable district-level routines. For teachers, the framework offers a developmental rather than punitive instrument; for institutions, it reconfigures the IQAC from a document collector into an analytical node that returns actionable feedback within fixed timelines; and for policymakers and accreditation agencies, it provides a replicable template for regional implementation that does not depend on metropolitan resource levels. IQAC coordinators, college principals, and accreditation bodies can begin adopting the three recommendations in their next annual cycle, with improved teaching quality expected ultimately to translate into measurable gains in student learning outcomes.

Several limitations qualify these claims and define the agenda for further research. The evidence, though gathered over an extended period from a principal site and several other institutions, remains concentrated within Assam, so the framework should be extended to other states and to comparable Global South settings to test whether the taxonomy travels intact or requires culturally local additions. The qualitative-descriptive design relies on student perceptions that, however well triangulated, cannot by themselves establish a causal link between teaching practice and learning outcomes, indicating a need for longitudinal mixed-methods studies that connect perceived teaching quality to measurable outcomes and graduate trajectories. Because the study includes no direct classroom observation, controlled implementations should test the proposed peer-review rubric and feedback timelines in practice. Finally, since both authors are faculty at the principal research site, insider bias was managed through reflexivity, member-checking, and independent coding but cannot be eliminated entirely, so replications by researchers external to the host institution would strengthen confidence in the transferability of both the taxonomy and the framework.

ACKNOWLEDGMENTS

The authors express their gratitude to all students, teachers, academic colleagues, and institutional authorities from various educational and research institutions of Assam who directly or indirectly supported this study through academic discussion, cooperation, and encouragement. The authors also sincerely acknowledge the cooperation of the participants involved in the research.

FUNDING INFORMATION

No funding is involved in this research study.

AUTHOR CONTRIBUTIONS STATEMENT

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

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Arabinda Rajkhowa	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
Munmi Dutta		✓				✓		✓		✓				

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

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no conflict of interest regarding the publication of this manuscript.

INFORMED CONSENT

Informed consent was obtained from all participants involved in the study. Participation was voluntary, and confidentiality was maintained throughout the research process.

ETHICAL APPROVAL

Prior institutional permission was obtained from North Lakhimpur University. The study complied with relevant ethical guidelines, and no personal identifiers were recorded.

DATA AVAILABILITY

The data supporting the findings of this study are available from the corresponding author, [AR], upon reasonable request, subject to confidentiality considerations.

REFERENCES

[1] F. Quansah, A. Cobbinah, K. Asamoah-Gyimah, and J. E. Hagan, "Validity of student evaluation of teaching in higher education: a systematic review," *Frontiers in Education*, vol. 9, p. 1329734, Feb. 2024, doi: 10.3389/feduc.2024.1329734.

[2] P. Spooren, B. Brockx, and D. Mortelmans, "On the validity of student evaluation of teaching," *Review of Educational Research*, vol. 83, no. 4, pp. 598–642, Dec. 2013, doi: 10.3102/0034654313496870.

[3] B. Uttl, C. A. White, and D. W. Gonzalez, "Meta-analysis of faculty's teaching effectiveness: student evaluation of teaching ratings and student learning are not related," *Studies in Educational Evaluation*, vol. 54, pp. 22–42, Sep. 2017, doi: 10.1016/j.stueduc.2016.08.007.

[4] B. Uttl, K. Cnudde, and C. A. White, "Conflict of interest explains the size of student evaluation of teaching and learning correlations in multisection studies: a meta-analysis," *PeerJ*, vol. 2019, no. 7, p. e7225, Jul. 2019, doi: 10.7717/peerj.7225.

[5] A. Boring, "Gender biases in student evaluations of teaching," *Journal of Public Economics*, vol. 145, pp. 27–41, Jan. 2017, doi: 10.1016/j.jpubeco.2016.11.006.

[6] T. Heffernan, "Sexism, racism, prejudice, and bias: a literature review and synthesis of research surrounding student evaluations of courses and teaching," *Assessment & Evaluation in Higher Education*, vol. 47, no. 1, pp. 144–154, Jan. 2022, doi: 10.1080/02602938.2021.1888075.

[7] F. Mengel, J. Sauermann, and U. Zölitz, "Gender bias in teaching evaluations," *Journal of the European Economic Association*, vol. 17, no. 2, pp. 535–566, Apr. 2019, doi: 10.1093/jeea/jvx057.

[8] T. Heffernan, "Abusive comments in student evaluations of courses and teaching: the attacks women and marginalised academics endure," *Higher Education*, vol. 85, no. 1, pp. 225–239, Jan. 2023, doi: 10.1007/s10734-022-00831-x.

[9] V. Hoorens, G. Dekkers, and E. Deschrijver, "Gender bias in student evaluations of teaching: students' self-affirmation reduces the bias by lowering evaluations of male professors," *Sex Roles*, vol. 84, no. 1–2, pp. 34–48, Jan. 2021, doi: 10.1007/s11199-020-01148-8.

[10] D. A. M. Peterson, L. A. Biederman, D. Andersen, T. M. Ditonto, and K. Roe, "Mitigating gender bias in student evaluations of teaching," *PLoS ONE*, vol. 14, no. 5, p. e0216241, May 2019, doi: 10.1371/journal.pone.0216241.

[11] K. Gelber, K. Brennan, D. Duriesmith, and E. Fenton, "Gendered mundanities: gender bias in student evaluations of teaching in political science," *Australian Journal of Political Science*, vol. 57, no. 2, p. 199, 2022, doi: 10.1080/10361146.2022.2043241.

[12] G. Hodson, "It is time to abandon student evaluations of teaching," *Nature Reviews Psychology*, vol. 4, no. 7, pp. 433–434, Apr. 2025, doi: 10.1038/s44159-025-00444-y.

[13] National Assessment and Accreditation Council, "Manual for self-study report: affiliated/constituent colleges," Bengaluru (Bangalore), 2019. [Online]. Available: https://www.naac.gov.in/

[14] Ministry of Education, Government of India, "National Education Policy 2020." Accessed: Apr. 27, 2026. [Online]. Available: https://www.education.gov.in/sites/upload_files/mhrd/files/NEP_Final_English_0.pdf

[15] University Grants Commission, "UGC Regulations on minimum qualifications for appointment of teachers and other academic staff in universities and colleges and measures for the maintenance of standards in higher education," New Delhi, 2018. [Online]. Available: https://www.ugc.gov.in/

[16] F. N.-A. Alhija, "Teaching in higher education: good teaching through students' lens," *Studies in Educational Evaluation*, vol. 54, pp. 4–12, Sep. 2017, doi: 10.1016/j.stueduc.2016.10.006.
[17] V. M. Catano and S. Harvey, "Student perception of teaching effectiveness: development and validation of the evaluation of teaching competencies scale (ETCS)," *Assessment & Evaluation in Higher Education*, vol. 36, no. 6, pp. 701–717, Oct. 2011, doi: 10.1080/02602938.2010.484879.
[18] M. Mimirinis and K. Ahlberg, "Variation in education doctoral students' conceptions of university teaching," *British Educational Research Journal*, vol. 47, no. 3, pp. 557–578, Jun. 2021, doi: 10.1002/berj.3669.
[19] P. Blumberg, "How critical reflection benefits faculty as they implement learner-centered teaching," *New Directions for Teaching and Learning*, vol. 2015, no. 144, pp. 87–97, Dec. 2015, doi: 10.1002/tl.20165.
[20] S. Myllykoski-Laine, A. Parpala, T. Hailikari, and L. Postareff, "Using self-reflection to support higher education teaching," *Reflective Practice*, vol. 25, no. 5, pp. 589–604, Sep. 2024, doi: 10.1080/14623943.2024.2376784.
[21] C. Constantinou and M. Wijnen-Meijer, "Student evaluations of teaching and the development of a comprehensive measure of teaching effectiveness for medical schools," *BMC Medical Education*, vol. 22, no. 1, p. 113, Feb. 2022, doi: 10.1186/s12909-022-03148-6.
[22] D. Clayson, "The student evaluation of teaching and likability: what the evaluations actually measure," *Assessment & Evaluation in Higher Education*, vol. 47, no. 2, pp. 313–326, Feb. 2022, doi: 10.1080/02602938.2021.1909702.
[23] L. Zhao, P. Xu, Y. Chen, and S. Yan, "A literature review of the research on students' evaluation of teaching in higher education," *Frontiers in Psychology*, vol. 13, p. 1004487, Sep. 2022, doi: 10.3389/fpsyg.2022.1004487.
[24] W. Stroebe, "Student evaluations of teaching encourages poor teaching and contributes to grade inflation: a theoretical and empirical analysis," *Basic and Applied Social Psychology*, vol. 42, no. 4, p. 276, 2020, doi: 10.1080/01973533.2020.1756817.
[25] J. D. Lilly, K. Wipawayangkool, and M. Pass, "Teaching evaluations and student grades: that's not fair!" *Journal of Management Education*, vol. 46, no. 6, pp. 994–1023, Dec. 2022, doi: 10.1177/10525629221084338.
[26] M. Witter and J. Hattie, "Can teacher quality be profiled? A cluster analysis of teachers' beliefs, practices and students' perceptions of effectiveness," *British Educational Research Journal*, vol. 50, no. 2, pp. 653–675, 2024, doi: 10.1002/berj.3938.
[27] A. Kulal, N. Abhishek, S. Dinesh, D. C. Bhat, and A. Girish, "Evaluating the promise and pitfalls of India's National Education Policy 2020: insights from the perspectives of students, teachers, and experts," *SAGE Open*, vol. 14, no. 4, pp. 1–18, Oct. 2024, doi: 10.1177/21582440241279367.
[28] K. Kishimoto, "Anti-racist pedagogy: from faculty's self-reflection to organizing within and beyond the classroom," *Race Ethnicity and Education*, vol. 21, no. 4, pp. 540–554, Jul. 2018, doi: 10.1080/13613324.2016.1248824.
[29] A. Segal, "Rethinking collective reflection in teacher professional development," *Journal of Teacher Education*, vol. 75, no. 2, pp. 155–167, Mar. 2024, doi: 10.1177/00224871231188702.
[30] H. Li, "Reflective practice for pre-service teachers' professional development," *SAGE Open*, vol. 15, no. 3, pp. 1–13, Jan. 2025, doi: 10.1177/21582440251363136.
[31] A. Karim, F. Amin, M. K. Kabilan, L. Reshmin, and E. U. Amin, "Pre-service TESOL teachers' professional-awareness building through reflecting on reflections of critical incidents," *Asian-Pacific Journal of Second and Foreign Language Education*, vol. 9, no. 1, p. 52, Apr. 2024, doi: 10.1186/s40862-024-00254-4.
[32] A. A. Juma, "Self-reflection in teaching: a comprehensive guide to empowering educators and enhancing student learning," *International Journal of Science and Research Archive*, vol. 12, no. 1, pp. 2835–2844, 2024, doi: 10.30574/ijsra.2024.12.1.1113.

BIOGRAPHIES OF AUTHORS

Arabinda Rajkhowa is an associate professor in the Department of Assamese, North Lakhimpur University, Assam, India. He holds a Ph.D. from Dibrugarh University and is a recognized research supervisor at both Dibrugarh University and North Lakhimpur University. Under his supervision, three scholars have been awarded Ph.D. degrees, while six research scholars are currently pursuing doctoral research. He has published over 30 research papers, and his research monographs have received notable scholarly recognition. He is a UGC Junior Research Fellowship recipient, has completed three UGC-funded research projects, and is a Sahitya Akademi Travel Grant awardee. His research interests encompass ecocriticism, Assamese literature, heritage studies, colonial socio-cultural studies, the environmental humanities, and the development of teaching attitudes in higher education. He can be contacted at email: arabinda@nlu.ac.in.

Munmi Dutta is a doctoral research scholar in the Department of Assamese at North Lakhimpur University. She holds an M.A. in Assamese from Dibrugarh University and is currently pursuing her Ph.D. research on the impact of modern science and technology on agrarian folk culture in Assam. She has presented more than 30 research papers at national and international conferences and has published 6 peer-reviewed research papers. Her research interests include the folk culture of Assam, agrarian traditions, postcolonial language and literature, medieval Assamese society and culture, and environmental conservation. In the present study, her contribution primarily relates to academic engagement with questions of teaching quality, pedagogy, and higher-education improvement. She can be contacted at email: munmi.dutta@nlu.ac.in.