

Assessment in the Crosshairs: How Contextual Factors Shape English Teachers' Conceptions of Assessment

Entika Fani Prastikawati ¹, Samantha Curle², Raveenthiran Vivekanantharasa³

¹Universitas Persatuan Guru Republik Indonesia Semarang, Indonesia

²University of Bath, United Kingdom

³The Open University of Sri Lanka, Sri Lanka

Abstract

This study explores English teachers' conceptions of assessment and the influence of contextual factors operating at the classroom levels within the Indonesian secondary education system. A representative survey was conducted among 236 English teachers from public and private secondary schools across Central Java, Indonesia. The findings reveal that teachers predominantly endorse assessment for purposes of student improvement, student accountability, and school accountability, while comparatively fewer teachers view assessment as irrelevant. Although many respondents express confidence in their assessment knowledge and perceived competence, concerns persist regarding the degree of professional trust they receive particularly in relation to assessment decision-making at both school and policy levels. To further explore underlying patterns, a latent class analysis (LCA) was conducted based on English teachers' responses to the "assessment as irrelevant" dimension. The analysis identified distinct groups of teachers, with those expressing higher levels of assessment irrelevance found across all schools and subject areas. These teachers tended to report lower confidence in assessment preparation, perceived less support from their institutions, and placed reduced value on academic achievement within societal contexts. The results underscore the complexity of teacher belief systems and the interplay of contextual forces shaping conceptions of assessment in Indonesia. Implications are discussed for assessment policy, school leadership, and professional development aimed at strengthening assessment literacy among English language teachers.

Keywords: conceptions of assessment; conceptual factors; English teachers

INTRODUCTION

In the era of rapid educational reform, assessment has become a pivotal force shaping curriculum implementation and teaching practices globally. Across diverse educational systems, there is a growing emphasis on aligning assessments with 21st-century skills such as critical thinking, creativity, and collaboration (Gan & Lam, 2024). However, this alignment cannot be achieved without understanding how teachers themselves conceptualize assessment, as their beliefs significantly influence assessment practices (Halim et al., 2024; Hung & Wu, 2024). In Indonesia, the Merdeka Curriculum reform advocates for more formative, student-centered approaches, aiming to move away from rigid summative testing structures. Yet, how English teachers respond to these shifts remains under-explored. Teachers' conceptions of assessment (COAs) reflect their beliefs about the purposes and roles of assessment ranging from improvement-oriented to accountability-focused or even viewing assessment as irrelevant (Brown et al., 2019; Xu et al., 2023). These conceptions often determine not just how assessment is conducted, but how learning itself is framed in the classroom. Given the socio-cultural complexity and hierarchical structure of Indonesia's education system, understanding these COAs is crucial. This study, therefore, investigates how contextual factors influence English teachers' conceptions of assessment in Indonesian schools.

Contextual factors can operate at multiple levels such as individual (micro), school (meso), and national or policy-related (macro), each exerting a distinct influence on teachers' beliefs and actions (Prastikawati et al., 2025; Williams-McBean, 2025). At the micro-level, individual teachers' training, experience, and confidence in assessment significantly shape their assessment practices. Research in various contexts, including Singapore and Spain, suggests that a teacher's self-perceived assessment literacy plays a strong role in whether they implement assessment as a tool for learning or compliance (Lin et al., 2017; Monteiro et al., 2021). At the meso level, school culture, leadership, and collegial support can either reinforce innovative assessment or perpetuate traditional exam-oriented systems

(Xu et al., 2023). In Indonesia, variations in school readiness and leadership commitment create inconsistencies in assessment implementation across schools (Mutiah, 2023). The macro-level, consisting of national policies and societal values, often sends mixed messages on the one hand promoting creativity and on the other reinforcing high-stakes testing through standardized exams (Brown et al., 2019). These contradictory pressures place Indonesian English teachers in a challenging position, affecting how they perceive and apply assessment in their classrooms. Therefore, examining these layers of influence offers insights into the complexities of assessment reform at the classroom level.

The conceptions of assessment held by teachers are far from homogenous. Brown and Remesal (2012) identified four major COAs: assessment for improvement, assessment for student accountability, assessment for school accountability, and assessment as irrelevant. Teachers may endorse multiple conceptions simultaneously, reflecting the multifaceted nature of their professional roles and contextual realities. For example, in high-stakes contexts, such as Indonesia's national exam system (*Ujian Nasional*), teachers may lean toward accountability-focused assessment (Suherman et al., 2025), while simultaneously believing in the pedagogical potential of formative feedback. This internal tension can result in inconsistencies between belief and practice, a phenomenon known as the values-practice gap (Ma & Bui, 2021). In some cases, teachers may view assessment as irrelevant due to limited professional development or insufficient institutional support (Xu & He, 2019). This diversity of views complicates top-down reform efforts that assume uniform teacher compliance or understanding. Recognizing the variability and the forces that shape it is thus a necessary step toward fostering more meaningful and sustainable assessment reforms. Investigating these conceptions among Indonesian English teachers allows for a nuanced understanding of how national directives interact with local beliefs and constraints.

Empirical studies have demonstrated that contextual factors do not influence teachers in isolation, but interact in complex and sometimes contradictory ways. For instance, Lin et al. (2017) in Singapore found that teachers' COAs ranged from conservative to progressive, shaped by both internal convictions and external pressures. In a similar vein, Fulmer et al. (2015) employed latent class analysis to reveal clusters of teachers with different assessment beliefs, irrespective of demographic categories such as experience or subject taught. These findings underscore the importance of examining beyond surface-level characteristics and looking into how teachers experience support, trust, and policy coherence. In Indonesia, research on teacher beliefs often focuses on general pedagogy, leaving assessment beliefs relatively under-theorized (Prastikawati et al., 2024). Moreover, school-level initiatives and teacher autonomy are often constrained by bureaucratic structures, making the enactment of formative and authentic assessment practices inconsistent (Baidoo-Anu et al., 2023). Teachers may express alignment with policy ideals but feel powerless to implement them due to lack of training, time, or institutional backing. This disconnect highlights the need for deeper inquiry into how Indonesian teachers navigate contextual forces in constructing their assessment beliefs and actions.

Although national curriculum policies, such as the *Merdeka* Curriculum, promote assessment for learning, such intentions may be undermined by entrenched educational traditions and systemic expectations. Indonesian schools remain heavily influenced by summative and norm-referenced assessment practices, where test scores are often tied to student promotion, teacher evaluation, and school ranking (Almossa & Alzahrani, 2022). These pressures reinforce a culture of accountability, often at the expense of formative, process-oriented assessment. In such a climate, teachers may struggle to reconcile their pedagogical ideals with institutional expectations, resulting in either resistance or superficial compliance. As Alonzo et al. (2021) noted, even when teachers report feeling competent in assessment, they may still lack trust, autonomy, or clarity from their schools and education systems. This phenomenon is mirrored in Indonesia, where the push for more innovative assessment has yet to be matched by coherent teacher training and structural support. Understanding the gap between policy discourse and teacher perception is critical to improving implementation fidelity and fostering authentic classroom assessment. Therefore, investigating how English teachers experience these tensions can illuminate both obstacles and entry points for change.

It is also worth considering the role of teacher agency in navigating contextual challenges. While systemic factors play a significant role, teachers are not passive recipients of policy but active agents who interpret, adapt, and sometimes resist imposed reforms (Hidri et al., 2024). Studies show

that teachers with high assessment literacy and professional confidence are more likely to implement innovative practices, even in rigid systems (Pardo et al., 2024). However, in Indonesia, limited access to professional learning communities and insufficient feedback mechanisms often hinder teachers' agency in assessment (Weng & Liu, 2024). School leaders also play a key role in either supporting or constraining teacher initiative, depending on how they interpret and enact policy at the school level. The interaction between personal beliefs, school culture, and policy expectations thus forms a complex ecology that shapes assessment practice. When this ecology is unsupportive or inconsistent, teachers may disengage or develop fragmented views of assessment. Consequently, mapping these experiences and perspectives offers valuable insights into the reform process and the conditions necessary for successful policy translation into practice.

Despite the increasing attention to educational assessment reforms in Indonesia, little is known about how English teachers construct their conceptions of assessment amid contextual constraints and policy expectations. Previous research often examines assessment knowledge or practice in isolation, without adequately capturing the belief systems that underlie teacher behavior. Furthermore, studies rarely disaggregate the multiple contextual layers such as individual, school, and national that influence these beliefs. As assessment continues to occupy a central role in Indonesia's educational transformation, understanding how teachers perceive and respond to this role becomes increasingly urgent. Doing so can inform more context-sensitive professional development, policy design, and school leadership strategies. This article seeks to contribute to this discourse by examining Indonesian English teachers' COAs through the lens of contextual influences, drawing parallels with international frameworks and findings. The study specifically addresses the gap between policy rhetoric and classroom realities as experienced by teachers on the ground. To that end, this study is guided by the following research questions: (1) What are Indonesian EFL teachers' conceptions of assessment? and (2) How do contextual factors shape Indonesian English teachers' conceptions of assessment?

METHODS

Respondents

This study employed a quantitative survey research design to investigate Indonesian secondary English teachers' conceptions of assessment and how these conceptions relate to contextual factors across micro-, meso-, and macro-levels. Following the conceptual framework proposed by Fulmer et al. (2015), which conceptualizes influences on teachers' assessment beliefs at the individual, school, and national levels, this study sought to identify patterns in teachers' beliefs and explore how they vary based on institutional and systemic influences. The study further utilized latent class analysis (LCA) to identify subgroups of teachers based on their endorsement of particular assessment beliefs, especially their views on assessment irrelevance. This design allowed for both descriptive and inferential examination of teacher belief patterns that might not be explained by demographic variables alone.

Participants and Context

The participants in this study comprised 236 English teachers from public and private secondary schools (junior and senior high schools) located across Central Java, Indonesia. Participants were selected through a two-stage sampling process. In the first stage, schools were selected using a stratified random sampling method to ensure representation from various geographic regions (urban, suburban, rural) and school types (public and private). In the second stage, English teachers from the selected schools were invited to participate in the study, either directly or through school-level coordinators. The final sample was broadly representative of the Central Java secondary school teacher population in terms of teaching experience, gender, and school affiliation. This region was selected due to its diversity in educational settings and its relevance in reflecting broader trends within the Indonesian education system.

Table 1 presents the demographic distribution of the 236 English teacher participants based on school type and geographic region within Central Java, Indonesia. Of the total sample, 162 teachers (68.6%) were from public secondary schools, while the remaining 74 teachers (31.4%) were affiliated with private institutions. In terms of geographic location, participants were drawn from schools

situated in urban, suburban, and rural areas to ensure broad representation of diverse teaching environments.

Table 1. Demographic Distribution of Participants by School Type and Geographic Region (N = 236)

Geographic Region	Public School	Private School	Total
Urban	54	28	82
Suburban	48	22	70
Rural	60	24	84
Total	162	74	236

Table 1 also presents the demographic distribution of the 236 English teacher participants based on school type and geographic region within Central Java, Indonesia. The participants were drawn from both public schools, which are government-funded institutions under the Ministry of Education, and private schools, which are independently managed, including both religious and non-religious institutions. In total, 162 teachers (68.6%) were affiliated with public secondary schools, while 74 teachers (31.4%) came from private schools. To ensure a broad representation of contextual teaching environments, participants were selected from three main geographic categories: urban, suburban, and rural areas. Urban schools refer to those located in provincial or major city centers, such as Semarang and Surakarta. In this category, 54 participants were from public schools and 28 from private schools, totaling 82 teachers (34.7%). Suburban schools are situated in satellite cities or towns adjacent to major urban centers; 48 teachers were from public suburban schools and 22 from private ones, making up 70 participants (29.7%). The rural category, encompassing schools in remote or less densely populated regions, had the highest representation with 84 teachers (35.6%), consisting of 60 from public schools and 24 from private schools.

The distribution across these categories reflects an intentional sampling approach that captures the diversity of school settings and institutional types in Central Java. This demographic composition allows for a meaningful investigation into how contextual factors such as school funding structure and geographic location may influence English teachers' conceptions of assessment. By including teachers from varied institutional contexts, the study enhances the interpretive depth and generalizability of its findings within the Indonesian secondary education landscape. This balanced yet diverse composition allowed for meaningful analysis of how contextual variables such as institutional setting and geographic location may influence teachers' conceptions of assessment. The inclusion of participants from a range of school types and locations also enhances the generalizability of the study's findings to the broader population of English teachers within the Central Java region.

Instruments

Three instruments were used in this study to gather data from participating teachers. The first instrument was an adaptation of Brown's (2006) Conceptions of Assessment (COA) inventory, consisting of 27 items rated on a six-point Likert scale (1 = strongly disagree to 6 = strongly agree). The instrument measured four key conceptions of assessment: student accountability, school accountability, student improvement, and assessment as irrelevant. The translated and culturally adapted version was validated through expert review and a pilot study conducted with 25 English teachers not included in the main sample. Internal consistency (Cronbach's alpha) was calculated for each subscale and yielded acceptable reliability: student accountability ($\alpha = .83$), school accountability ($\alpha = .78$), improvement ($\alpha = .85$), and irrelevance ($\alpha = .80$).

The second instrument measured contextual factors affecting assessment beliefs, developed with reference to Fulmer et al. (2015). This instrument consisted of 29 items rated on a four-point Likert scale (1 = strongly disagree to 4 = strongly agree). The items addressed micro-level (e.g., teacher competence, emotional factors), meso-level (e.g., school leadership, peer support), and macro-level (e.g., national policy clarity, societal expectations) influences. While the instrument was not structured as distinct subscales, items were categorized for analysis purposes based on their alignment with the three contextual levels. Furthermore, the third instrument collected teacher demographic information, including gender, total years of teaching experience, years at current school, highest qualification,

school type (public or private), and school location (urban or rural). These variables provided additional background for understanding variation in conceptions of assessment.

Data collection procedures

Data were collected using an online survey platform, which allowed for efficient distribution and tracking of responses across geographically dispersed schools. Prior to full deployment, a pilot test of the survey instruments was conducted to ensure clarity and cultural appropriateness of the items. Teachers were invited to participate via official school email or through direct contact with school administrators and English teacher associations. Participation was voluntary, and informed consent was obtained from all respondents. The survey was open for four weeks, with periodic reminders sent to improve response rates. At the close of the data collection period, 236 complete and valid responses were retained for analysis.

To preserve confidentiality, individual identities and school affiliations were anonymized. Ethical clearance for the study was obtained from the institutional research ethics committee, and permissions were secured from the relevant regional education authorities in Central Java.

Data analysis

Data analysis was conducted in two stages. In the first stage, descriptive statistics (means, standard deviations) and bivariate correlations were computed to summarize teachers' conceptions of assessment and their reported contextual perceptions. Analysis of variance (ANOVA) and multiple regression were conducted to examine relationships between COA subscales and demographic or contextual variables. Reliability indices (Cronbach's alpha) were used to assess internal consistency of the COA subscales.

In the second stage, latent class analysis (LCA) was applied using the *poLCA* package in R (Linzer & Lewis, 2011) to explore hidden patterns in teachers' responses to the assessment as irrelevant subscale. The goal of this analysis was to identify distinct groups of teachers whose beliefs diverged from expected demographic patterns. Models were estimated for one to six latent groups, and the best-fitting model was selected based on Bayesian Information Criterion (BIC) and Akaike's Information Criterion with correction (AICc) values, with lower values indicating better model fit (Burnham & Anderson, 2004). Each teacher was assigned to the latent group to which they had the highest probability of membership. Subsequent ANOVA tests compared latent groups across contextual factor items to explore whether patterns of assessment irrelevance were associated with differences in school support, teacher preparation, or societal values. This combination of psychometric, correlational, and classification-based analyses provided a robust framework for interpreting the complex relationships among Indonesian English teachers' assessment beliefs and their professional contexts.

FINDINGS AND DISCUSSION

Understanding teachers' conceptions of assessment is fundamental to advancing assessment reform, particularly in systems navigating the balance between traditional, high-stakes testing and more formative, learner-centered approaches. Conceptions of assessment is defined as teachers' beliefs about the purposes, functions, and value of assessment play a pivotal role in shaping instructional decisions and classroom practices (Prastikawati et al., 2022). In the Indonesian context, where recent curriculum reforms such as the Merdeka Curriculum emphasize assessment for learning, exploring how teachers perceive assessment is both timely and necessary. Prior research has shown that teachers often hold multiple, and sometimes conflicting, beliefs about assessment, including its use for student improvement, accountability, and, in some cases, as an irrelevant or burdensome task (Fulmer et al., 2015; Lin et al., 2017). To examine these dimensions, this study utilized Conceptions of Assessment framework (Brown et al., 2019) and collected data from a representative sample of secondary school English teachers across Central Java, Indonesia. The following section presents the descriptive statistics and intercorrelations among the four core COA dimensions namely student accountability, school accountability, student improvement, and assessment as irrelevant, as reported by the participating teachers. These results provide insight into the dominant assessment beliefs among Indonesian English teachers and serve as the foundation for deeper analyses of contextual influences

Table 2 presents the descriptive statistics and intercorrelations among the four subscales of the Conceptions of Assessment (COA) instrument used in this study: Student Accountability, School Accountability, Student Improvement, and Assessment as Irrelevant. These subscales reflect teachers' beliefs regarding the roles and purposes of assessment in educational contexts, as originally developed by Brown et al. (2019). The means and standard deviations (SD) indicate general trends in the teachers' agreement with each conception, while Pearson's r values reveal the strength and direction of the relationships between subscales.

Conceptions of Assessment among Indonesian EFL Teachers

The data show that English teachers in this sample generally agreed with the constructive roles of assessment. The highest mean score was observed for the Student Improvement subscale ($M = 4.62$, $SD = 0.58$), suggesting that most teachers viewed assessment as a tool to enhance student learning and development. This was followed by School Accountability ($M = 4.45$, $SD = 0.78$) and Student Accountability ($M = 4.20$, $SD = 0.85$), indicating moderate support for the belief that assessment holds schools and students accountable for performance outcomes. In contrast, the lowest mean score was found for the Assessment as Irrelevant subscale ($M = 3.12$, $SD = 0.79$), indicating a general disagreement with the idea that assessment lacks educational value.

Table 2. Result of English teachers' COA

COA Subscale	Mean	SD	(1)	(2)	(3)	(4)
Student Accountability	4.20	0.85	–			
School Accountability	4.45	0.78	.31***	–		
Student Improvement	4.62	0.58	.42***	.29***	–	
Assessment as Irrelevant	3.12	0.79	-.05	-.01	-.18**	–

The correlation analysis reveals several statistically significant and positive associations among the constructive COA subscales. A moderate positive correlation was found between Student Accountability and Student Improvement ($r = .42$, $p < .001$), and between Student Accountability and School Accountability ($r = .31$, $p < .001$). These relationships suggest that teachers who believe assessment is useful for holding students accountable also tend to value it for both improving learning and institutional accountability. Additionally, School Accountability was positively associated with Student Improvement ($r = .29$, $p < .001$), reinforcing the interconnected nature of teachers' constructive assessment conceptions.

In contrast, the Assessment as Irrelevant subscale showed weak or negative relationships with the other three subscales. A small but significant negative correlation was observed between Assessment as Irrelevant and Student Improvement ($r = -.18$, $p < .01$), indicating that teachers who viewed assessment as irrelevant were less likely to see it as beneficial for improving student learning. No significant correlations were found between Assessment as Irrelevant and either Student Accountability ($r = -.05$, n.s.) or School Accountability ($r = -.01$, n.s.), suggesting that teachers' endorsement of the irrelevance of assessment is relatively independent of their accountability beliefs.

These findings are largely consistent with prior research (Fulmer et al., 2015), which demonstrated that conceptions of assessment for improvement and accountability tend to be interrelated, whereas the belief in assessment's irrelevance functions as a distinct dimension. The presence of weak or no associations between irrelevance and other subscales may indicate a unique profile of teachers who feel disempowered, unsupported, or disconnected from institutional assessment goals. These individuals warrant further attention in assessment reform efforts, particularly in contexts such as Indonesia, where tensions between top-down accountability pressures and bottom-up classroom realities continue to shape teachers' beliefs and practices.

Table 3. Result of English teachers' COA

Variable	Group A (Low Irrelevance, n = 144)	Group B (High Irrelevance, n = 92)
Student Accountability	4.35 (0.76)*	4.05 (0.89)*
School Accountability	4.52 (0.69)	4.38 (0.73)
Student Improvement	4.71 (0.53)*	4.41 (0.60)*
Assessment as Irrelevant	2.78 (0.55)*	4.01 (0.51)*

Total Years of Teaching Experience	12.6 (9.1)	12.1 (8.8)
Years at Current School	7.2 (6.5)	7.0 (6.7)

Furthermore, table 3 describes the result of LCA. Table 3 revealed two distinct groups of teachers based on their endorsement of the belief that assessment is irrelevant. To further explore variations in teachers' conceptions of assessment, particularly regarding their endorsement of assessment as irrelevant, a latent class analysis (LCA) was conducted using teachers' responses to the Assessment as Irrelevant subscale. This analytical approach allows for the identification of unobserved (latent) subgroups within the population based on patterns in their item-level responses, rather than relying on demographic characteristics alone. Models with one to six latent groups were estimated using the poLCA package in R (Linzer & Lewis, 2011), and model fit was evaluated using the Bayesian Information Criterion (BIC) and the corrected Akaike Information Criterion (AICc), both of which prioritize model parsimony and fit, with lower values indicating better-fitting models (Burnham & Anderson, 2004).

Among the estimated models, the two-group solution demonstrated the optimal fit, with BIC = 5531.21 and AICc = 5374.08, and was therefore selected for further interpretation. Each teacher was assigned to one of the two latent groups based on their highest posterior membership probability. The results, displayed in Table 3, indicate meaningful differences between the two groups in their conceptions of assessment. Group A (n = 144) consisted of teachers who disagreed with the notion that assessment is irrelevant. This group reported higher mean scores on the student accountability (M = 4.35, SD = 0.76), school accountability (M = 4.52, SD = 0.69), and student improvement (M = 4.71, SD = 0.53) subscales. Conversely, Group B (n = 92) represented teachers who agreed more strongly with the idea that assessment is irrelevant (M = 4.01, SD = 0.51), and demonstrated lower agreement with assessment for student accountability (M = 4.05, SD = 0.89) and student improvement (M = 4.41, SD = 0.60). The differences between the two groups on these subscales were statistically significant ($p < .05$), suggesting distinct conceptions of assessment between groups.

Interestingly, there were no significant differences between Group A and Group B in terms of demographic characteristics, such as total years of teaching experience (Group A: M = 12.6 years, SD = 9.1; Group B: M = 12.1 years, SD = 8.8) or years teaching at the current school (Group A: M = 7.2 years, SD = 6.5; Group B: M = 7.0 years, SD = 6.7). This suggests that endorsement of assessment as irrelevant is not strongly linked to conventional background variables, but instead may reflect deeper, contextually driven beliefs related to assessment preparation, professional trust, or institutional support.

These findings reinforce earlier work (DeLuca et al., 2021; Fulmer et al., 2015) that emphasizes the importance of examining latent belief structures, rather than assuming uniformity across teacher populations. The existence of a sizeable group of teachers who view assessment as irrelevant, despite professional experience, highlights the need for targeted interventions aimed at assessment literacy and institutional culture reform.

Contextual Factors Shaping Indonesian English Teachers' Conceptions of Assessment

These findings reinforce earlier work (Atjonen et al., 2022) that emphasizes the importance of examining latent belief structures, rather than assuming uniformity across teacher populations.

Table 1. Contextual factors by latent group A and group B

Contextual Factors	Group A (Low Irrelevance, n = 144)	Group B (High Irrelevance, n = 92)	Mean Difference
Micro Level			
Assessment Competence	3.51 (0.42)*	3.22 (0.49)*	.29*
Emotional Confidence	3.40 (0.41)*	3.28 (0.46)*	.12*
Meso Level			
School Leadership Support	3.36 (0.50)*	3.05 (0.53)*	.31*
Peer Collaboration	3.27 (0.47)	3.11 (0.49)	.16
Macro Level			
Policy Clarity	3.05 (0.58)*	2.83 (0.63)*	.22*
Societal Expectations of Success	3.19 (0.54)	3.09 (0.60)	.10

Table 4 displays the average scores of contextual factors at the micro, meso, and macro levels for the two latent groups discovered in the preceding analysis: Group A (teachers with low endorsement of assessment irrelevance) and Group B (teachers with strong endorsement of assessment irrelevance). The findings reveal distinct and consistent disparities in the perceptions of different groups regarding their professional and institutional contexts. At the micro level, Group A exhibited much superior assessment competence ($M = 3.51$, $SD = 0.42$) relative to Group B ($M = 3.22$, $SD = 0.49$), resulting in a mean difference of .29 ($p < .05$). Teachers in Group A had higher emotional confidence in implementing assessment techniques ($M = 3.40$, $SD = 0.41$) compared to their counterparts in Group B ($M = 3.28$, $SD = 0.46$). These findings indicate that educators who dismiss the notion of assessment as inconsequential are more inclined to perceive themselves as competent and emotionally stable in their assessment responsibilities, consistent with other research that associates self-efficacy with positive assessment beliefs (van der Linden et al., 2023).

At the meso level, notable group disparities were also identified. Educators in Group A evaluated their institutions as offering superior leadership support for assessment ($M = 3.36$, $SD = 0.50$) in contrast to those in Group B ($M = 3.05$, $SD = 0.53$), resulting in a mean difference of 0.31 ($p < .05$). Despite the lack of statistical significance in peer collaboration differences, a consistent pattern emerged, with Group A teachers ($M = 3.27$, $SD = 0.47$) reporting greater collegial support than Group B teachers ($M = 3.11$, $SD = 0.49$). The findings indicate that supportive school cultures, particularly through leadership involvement, may enhance teachers' positive perceptions regarding evaluation.

At the macro level, disparities across the groups were most pronounced regarding policy clarity. Teachers in Group A expressed a higher level of consensus regarding the clarity and coherence of national assessment policies ($M = 3.05$, $SD = 0.58$) compared to teachers in Group B ($M = 2.83$, $SD = 0.63$), resulting in a mean difference of 0.22 ($p < .05$). No substantial differences between groups were noted regarding societal expectations of academic success, although Group A's mean ($M = 3.19$, $SD = 0.54$) was marginally higher than that of Group B ($M = 3.09$, $SD = 0.60$). This conclusion suggests that although both groups acknowledge the significant influence of societal influences on assessment methods in Indonesia, it is the clarity of regulation that most distinctly differentiates teachers' perceptions of assessment.

The results underscore the importance of contextual support in influencing the evaluation beliefs of Indonesian English teachers. Educators who regard themselves as proficient, bolstered by their institutions, and directed by explicit policy frameworks are less inclined to consider assessment as inconsequential. In contrast, diminished feelings of support and systemic clarity seem to foster disengagement from the instructional value of evaluation. The findings correspond with global data (Weng & Liu, 2024), underscoring the necessity of synchronizing teacher preparation, school leadership, and policy communication to promote positive perceptions of assessment.

Discussion

The findings of this study provide valuable insights into Indonesian English teachers' conceptions of assessment and the role of contextual factors in shaping these beliefs. Overall, teachers demonstrated moderate to strong endorsement of assessment for student accountability, school accountability, and student improvement, while showing weaker endorsement of assessment as irrelevant. This pattern aligns with previous research by Yildirim et al. (2024), which demonstrated that teachers often hold constructive views of assessment but may vary in the extent to which they dismiss it as an irrelevant activity. Importantly, the Indonesian context reflects a transitional stage in curriculum reform, particularly with the implementation of the Merdeka Curriculum, which emphasizes assessment for learning rather than high-stakes testing. Teachers' beliefs about assessment thus become critical indicators of reform readiness, as misalignments between teacher conceptions and policy intentions can hinder implementation. These results contribute to the broader literature on assessment literacy, particularly in developing contexts where educational reforms often collide with entrenched assessment traditions.

The latent class analysis (LCA) provided a deeper understanding of the heterogeneity among teachers by identifying two distinct groups based on their endorsement of assessment irrelevance. Group A, representing teachers with low endorsement of irrelevance, tended to strongly support assessment for improvement and accountability purposes. Group B, in contrast, showed higher

endorsement of irrelevance and weaker support for constructive conceptions. Notably, demographic factors such as teaching experience and years in the current school did not significantly distinguish these groups, indicating that professional tenure alone does not explain differences in beliefs. This finding mirrors the results of (Fulmer et al., 2015; Muianga, 2023), who also observed that assessment conceptions are shaped more by contextual and institutional influences than by teacher background characteristics. These insights suggest that interventions to enhance assessment literacy must prioritize contextual supports rather than assuming that experience alone leads to positive beliefs. The existence of a sizeable group of teachers who perceive assessment as irrelevant underscores the urgent need for targeted professional development that fosters constructive engagement with assessment.

At the micro level, teachers' self-perceived competence and emotional confidence in assessment emerged as significant predictors of their conceptions. Teachers in Group A reported higher confidence and competence compared to their counterparts in Group B, reinforcing the idea that assessment self-efficacy is closely linked to positive beliefs about assessment. This finding resonates with earlier studies by (Aksu & Gungor, 2025), who argued that teachers' self-efficacy mediates the relationship between their assessment literacy and classroom practices. It also highlights the psychological dimension of assessment, suggesting that teachers' confidence in their ability to design, implement, and interpret assessment tasks can directly influence whether they perceive assessment as a valuable or burdensome activity. In the Indonesian context, where teachers are expected to integrate formative assessment within the Merdeka Curriculum, enhancing teachers' sense of competence is crucial. Without adequate support, teachers may fall back on summative practices or disengage from assessment entirely, reinforcing perceptions of irrelevance. Thus, professional training programs should not only build technical skills but also cultivate teachers' confidence in their assessment capacities.

At the meso level, school leadership support and peer collaboration emerged as important factors differentiating the two latent groups. Teachers in Group A perceived higher levels of school support, particularly from principals and leadership teams, than those in Group B. This finding supports the literature emphasizing the centrality of school culture and leadership in shaping teachers' beliefs and practices (Xu & He, 2019). In schools where leaders actively promote a culture of assessment for learning, teachers are more likely to embrace constructive views of assessment. Conversely, limited leadership engagement and insufficient collegial collaboration can reinforce negative perceptions and disengagement. The Indonesian educational system, with its growing emphasis on school autonomy under the curriculum, makes the role of leadership even more critical. School leaders are expected not only to ensure compliance with policy but also to foster environments where teachers feel supported in implementing innovative assessment practices. Strengthening leadership training in assessment literacy is therefore a promising avenue to address the gap between policy aspirations and classroom realities.

At the macro level, policy clarity emerged as a distinguishing factor between the groups, whereas societal expectations of academic success did not show significant differences. Teachers in Group A reported perceiving national assessment policies as clearer and more coherent, which was associated with more constructive conceptions of assessment. This finding is consistent with (Deneen & Brown, 2016), who noted that unclear or inconsistent policy messages undermine teacher trust in assessment systems. In the Indonesian case, where significant reforms such as the replacement of high-stakes examinations with holistic evaluations are being introduced, policy clarity becomes especially critical. Teachers who perceive the reforms as ambiguous may be more inclined to dismiss assessment as irrelevant, regardless of their experience or school support. The limited role of societal expectations as a differentiator suggests that while cultural pressures for academic achievement remain strong in Indonesia, they are less influential than institutional and systemic clarity in shaping teachers' beliefs. This highlights the importance of strengthening communication strategies at the national level to ensure teachers interpret reforms as coherent, realistic, and supportive of their classroom practice.

Taken together, these findings highlight the complex interplay of micro, meso, and macro contextual factors in shaping teachers' conceptions of assessment. The results reinforce the need for a multi-level approach to strengthening assessment literacy and fostering constructive beliefs among teachers. At the micro level, professional development must emphasize both competence and confidence in assessment practices. At the meso level, leadership and collegial support should be prioritized to create school cultures that value assessment for learning. At the macro level,

policymakers must ensure clarity, consistency, and practicality in assessment reforms, particularly in contexts undergoing rapid curricular transformation such as Indonesia. These results contribute to the growing body of literature demonstrating that teachers' conceptions of assessment are not isolated cognitive beliefs but are deeply embedded within institutional and systemic environments (Baidoo-Anu et al., 2023; Ma & Bui, 2021). For Indonesia, the successful implementation of the Merdeka Curriculum depends not only on policy design but also on aligning contextual supports to cultivate teachers' constructive engagement with assessment.

CONCLUSION

This study examined Indonesian English teachers' conceptions of assessment and the contextual factors shaping these beliefs at the micro, meso, and macro levels. The findings revealed that while teachers generally endorsed constructive purposes of assessment such as accountability and student improvement, a notable subgroup perceived assessment as irrelevant. The latent class analysis highlighted that differences between groups were not explained by demographic characteristics such as teaching experience, but rather by contextual factors including teachers' sense of competence, school leadership support, and perceptions of policy clarity. These results underscore the multifaceted nature of assessment conceptions and the importance of institutional and systemic conditions in shaping teachers' beliefs.

The study contributes to the growing literature on assessment literacy by demonstrating how contextual environments influence teachers' assessment conceptions in the Indonesian context. At the micro level, professional competence and confidence emerged as critical, suggesting that teacher preparation programs and in-service training should go beyond technical knowledge to build confidence in assessment practices. At the meso level, school leaders play a central role in shaping assessment culture, indicating that leadership training should explicitly include assessment literacy. At the macro level, policy clarity was shown to be decisive, highlighting the need for transparent and consistent communication of reforms such as those embedded in the curriculum. Together, these implications suggest that successful implementation of assessment reform requires a systemic approach that integrates teacher capacity-building, school leadership, and coherent policy frameworks.

Future research should build on these findings by exploring the longitudinal development of teachers' conceptions of assessment over time, particularly as Indonesia continues to implement the current curriculum. Longitudinal studies could capture how shifts in policy, professional development initiatives, and school environments influence changes in teachers' beliefs and practices. Additionally, mixed-methods research combining surveys with classroom observations and interviews would provide deeper insights into how teachers' stated beliefs align or diverge from their enacted assessment practices. Comparative studies across different provinces or countries could also highlight how cultural and institutional contexts shape similarities and differences in assessment conceptions.

Finally, this study emphasizes that teachers' conceptions of assessment cannot be fully understood in isolation from their contextual realities. Strengthening assessment literacy in Indonesia will require coordinated efforts that recognize the interdependence of micro, meso, and macro factors. By fostering teacher competence, enhancing school leadership support, and ensuring coherent and transparent policies, stakeholders can cultivate more constructive and sustainable assessment practices. Such efforts are not only essential for the effective implementation of the Merdeka Curriculum but also for advancing equitable and meaningful learning opportunities for Indonesian students in an evolving educational landscape.

FUNDING STATEMENT

This study does not receive any fundings from any agencies or institutions.

REFERENCES

Abd Halim, H., Hamzah, M. I., & Zulkifli, H. (2024). A systematic review on the formative assessment practice in teaching and learning in secondary school. *International Journal of*

- Evaluation and Research in Education*, 13(2), 1173-1183.
<https://doi.org/10.11591/ijere.v13i2.26187>
- Aksu, İ., & Gungor, M. N. (2025). The Interplay between Assessment Conceptions and Practices of EFL Instructors: A Mixed-method Study. *Kastamonu Education Journal*, 33(3), 646-657.
<https://doi.org/10.24106/kefdergi.1750257>
- Almossa, S. Y., & Alzahrani, S. M. (2022). Assessment practices in Saudi higher education during the COVID-19 pandemic. *Humanities and Social Sciences Communications*, 9(1).
<https://doi.org/10.1057/s41599-021-01025-z>
- Alonzo, D., Labad, V., Bejano, J., & Guerra, F. (2021). The policy-driven dimensions of teacher beliefs about assessment. *Australian Journal of Teacher Education (Online)*, 46(3), 36-52.
<https://doi.org/10.14221/ajte.2021v46n3.3>
- Atjonen, P., Pöntinen, S., Kontkanen, S., & Ruotsalainen, P. (2022, July). In enhancing preservice teachers' assessment literacy: Focus on knowledge base, conceptions of assessment, and teacher learning. In *Frontiers in Education (Vol. 7, p. 891391)*. *Frontiers Media SA*.
<https://doi.org/10.3389/educ.2022.891391>
- Baidoo-Anu, D., Rasooli, A., DeLuca, C., & Cheng, L. (2023). Conceptions of classroom assessment and approaches to grading: Teachers' and students' perspectives. *Education Inquiry*, 1-29.
<https://doi.org/10.1080/20004508.2023.2244136>
- Brown, G. T., & Remesal, A. (2012). Prospective teachers' conceptions of assessment: A cross-cultural comparison. *The Spanish journal of psychology*, 15(1), 75-89.
https://doi.org/10.5209/rev_SJOP.2012.v15.n1.37286
- Brown, G. T., Gebril, A., & Michaelides, M. P. (2019, March). Teachers' conceptions of assessment: A global phenomenon or a global localism. In *Frontiers in education (Vol. 4, p. 16)*. *Frontiers Media SA*. <https://doi.org/10.3389/educ.2019.00016>
- Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Methods & Research*, 33, 261-304.
<https://doi.org/10.1177/0049124104268644>
- DeLuca, C., Rickey, N., & Coombs, A. (2021). Exploring assessment across cultures: Teachers' approaches to assessment in the US, China, and Canada. *Cogent Education*, 8(1), 1921903.
<https://doi.org/10.1080/2331186X.2021.1921903>
- Deneen, C. C., & Brown, G. T. (2016). The impact of conceptions of assessment on assessment literacy in a teacher education program. *Cogent Education*, 3(1), 1225380.
<https://doi.org/10.1080/2331186X.2016.1225380>
- Fulmer, G. W., Lee, I. C., & Tan, K. H. (2015). Multi-level model of contextual factors and teachers' assessment practices: An integrative review of research. *Assessment in Education: Principles, Policy & Practice*, 22(4), 475-494. <https://doi.org/10.1080/0969594X.2015.1017445>
- Gan, L., & Lam, R. (2024). Unveiling Classroom Assessment Literacy: Does Teachers' Self-Directed Development Play Out?. *Education Sciences*, 14(9), 961.
<https://doi.org/10.3390/educsci14090961>
- Hidri, S., Aziz, M., & Qutub, M. (2024). Conceptions of Assessment Among Pakistani Teachers of English: Implications for Policy and Professional Development. *Heliyon*, 10(17).
<https://doi.org/10.1016/j.heliyon.2024.e36780>
- Hung, C. S., & Wu, H. K. (2024). High school science teachers' assessment literacy for inquiry-based science instruction. *International Journal of Science Education*, 46(7), 621-642.
<https://doi.org/10.1080/09500693.2023.2251657>
- Linzer, D. A., & Lewis, J. (2011). poLCA: An R package for polytomous variable latent class analysis. *Journal of Statistical Software*, 42, 1-29. <https://doi.org/10.18637/jss.v042.i10>
- Lin, T. J., Tan, A. L., Lee, M. H., & Tsai, C. C. (2017). Evaluating and comparing Singaporean and Taiwanese eighth graders' conceptions of science assessment. *Research in Science & Technological Education*, 35(4), 391-408. <https://doi.org/10.1080/02635143.2017.1318363>
- Ma, M., & Bui, G. (2021). Chinese secondary school teachers' conceptions of L2 assessment: A mixed-methods study. *Studies in Second Language Learning and Teaching*, 11(3), 445-472.
<https://doi.org/10.14746/ssllt.2021.11.3.7>

- Monteiro, V., Mata, L., & Santos, N. N. (2021, April). Assessment conceptions and practices: Perspectives of primary school teachers and students. In *Frontiers in Education* (Vol. 6, p. 631185). Frontiers Media SA. <https://doi.org/10.3389/feduc.2021.631185>
- Muianga, F. (2023, January). English language teachers' conceptions of assessment. In *Frontiers in Education* (Vol. 7, p. 972005). Frontiers Media SA. <https://doi.org/10.3389/feduc.2022.972005>
- Mutiah, S. D. (2023). Embedding Formative Assessment: A Framework of Strategies in Formative Assessment. *International Journal of Education, Information Technology, and Others*, 6(2), 372-382.
- Pardo, R., García-Pérez, D., & Panadero, E. (2024). Shaping the assessors of tomorrow: How practicum experiences develop assessment literacy in secondary education pre-service teachers. *Teaching and Teacher Education*, 152, 104798. <https://doi.org/10.1016/j.tate.2024.104798>
- Prastikawati, E. F., Mujiyanto, J., Saleh, M., & WuliFitriati, S. (2022). Pre-service EFL Teachers' Conceptions of Assessment During Their Teaching Practicum. *KnE Social Sciences*, 615-626. <https://doi.org/10.18502/kss.v7i19.12480>
- Prastikawati, E. F., Adeoye, M. A., & Ryan, J. C. (2024). Fostering effective teaching practices: Integrating formative assessment and mentorship in Indonesian preservice teacher education. *Indonesian Journal on Learning and Advanced Education (IJOLAE)*, 230-253. <https://doi.org/10.23917/ijolae.v6i2.23431>
- Prastikawati, E. F., Mujiyanto, J., Saleh, M., & Fitriati, S. W. (2025). Factors Influencing Assessment Decision-Making of Pre-Service EFL Teachers: Insights From the Teaching Practicum Experience. *Theory and Practice in Language Studies*, 15(4), 1330-1339. <https://doi.org/10.17507/tpls.1504.31>
- Suherman, A., Surayya, S. A., Dwimaulani, A., & Setiawan, B. (2025). A Study of Formative Assessment Literacy in Indonesia: Insight from Pre-Service EFL Teachers. *Scope: Journal of English Language Teaching*, 9(2), 856-866. <https://doi.org/10.30998/scope.v9i2.26043>
- van der Linden, J., van der Vleuten, C., Nieuwenhuis, L., & van Schilt-Mol, T. (2023). Formative use of assessment to foster self-regulated learning: the alignment of teachers' conceptions and classroom assessment practices. *Journal of Formative Design in Learning*, 7(2), 195-207. <https://doi.org/10.1007/s41686-023-00082-8>
- Weng, F., & Liu, X. (2024). Exploring second language students' language assessment literacy: impact on test anxiety and motivation. *Frontiers in psychology*, 15, 1289126. <https://doi.org/10.3389/fpsyg.2024.1289126>
- Williams-McBean, C. (2025). Factors influencing teachers' choice and use of assessment. *International Journal of Studies in Education and Science*, 6(2), 212-239. <https://doi.org/10.46328/ijses.129>
- Xu, Y., & He, L. (2019, December). How pre-service teachers' conceptions of assessment change over practicum: Implications for teacher assessment literacy. In *Frontiers in Education* (Vol. 4, p. 145). Frontiers Media SA. <https://doi.org/10.3389/feduc.2019.00145>
- Xu, Y., Chen, J., & He, L. (2023). How personal, experiential, and contextual factors mediate EFL teachers' conceptions of assessment: A narrative study. *Chinese Journal of Applied Linguistics*, 46(2), 237-252. <https://doi.org/10.1515/CJAL-2023-0206>
- Yildirim, A., Oscarson, A. D., Hilden, R., & Fröjdendahl, B. (2024). Teaching summative assessment: A curriculum analysis of pre-service language teacher education in Sweden and Finland. *Journal of Teacher Education*, 75(2), 203-218. <https://doi.org/10.1177/00224871231214799>