



Language Assessment in Multicultural Primary Classrooms in Denpasar: Teachers' Challenges and Strategies

Kadek Dessy Desmiyani¹, Ni Luh Putu Eka Sulistia Dewi²,
Ni Putu Era Marsakawati³

^{1,2,3}English Education, Ganesha University of Education, Bali

Article Info	Abstract
Received: 2026-06-22 Revised: 2026-08-15 Accepted: 2026-09-21 Keywords: <i>challenges and strategies; language assessment; multicultural classroom; responsive language assessment</i> DOI: 10.24256/ideas.v14i2.11387 Corresponding Author: Kadek Dessy Desmiyani dessy.desmiyani@student.undiksha.ac.id Ganesha University of Education, Bali	<i>Language assessment in multicultural primary classrooms requires teachers to distinguish linguistic difficulty from content knowledge while responding to culturally shaped ways of participating and demonstrating learning. This study investigates the challenges teachers encounter in conducting language assessment in multicultural primary schools in Denpasar, Bali, and the strategies they use to make assessment more equitable and meaningful. Using a qualitative case study design, the study involved 12 teachers from public and private primary schools selected for their experience in multicultural classrooms. Data were generated through semi-structured interviews, classroom observations, and document analysis and were analysed thematically following Braun and Clarke's procedures of familiarisation, coding, theme development, review, and naming. The findings show three interrelated areas of challenge: linguistic and comprehension barriers, cultural misalignment in interpreting student participation and performance, and limited access to culturally responsive assessment resources and institutional support. Teachers responded through differentiated assessment formats, visual and linguistic scaffolding, translanguaging, culturally relevant tasks, and regular formative feedback. The study is novel in examining teacher-reported challenges and responsive assessment strategies within the same multicultural primary-school cases in Denpasar, an urban context shaped by local Balinese culture, inter-island mobility, and international migration. The findings indicate that equitable language assessment depends not only on technical assessment quality but also on cultural validity, flexible opportunities for expression, and sustained institutional support. These context-specific findings extend discussion of culturally and linguistically responsive assessment in multilingual primary education.</i>

1. Introduction

Language assessment is a central component of primary language education because it provides evidence about learners' development in listening, speaking, reading, and writing and informs subsequent instructional decisions. Classical assessment principles emphasise practicality, reliability, validity, authenticity, and washback (Bachman & Palmer, 1996; Brown, 2004). In multicultural classrooms, however, these principles cannot be interpreted only as technical properties of tests. Assessment is also a social practice in which language, culture, identity, and classroom expectations shape how learners understand tasks and how teachers interpret performance (McNamara, 2001).

Cultural and linguistic diversity complicates assessment because students may differ substantially in proficiency in the language of schooling, familiarity with task formats, participation norms, and opportunities to use their full linguistic repertoires. Recent research has shown that assessments that overlook these differences may underrepresent multilingual learners' knowledge. Rosheim et al. (2024), for example, found that literacy assessments could reproduce inequities when texts lacked linguistic and cultural relevance, while Yang (2025) showed that linguistically responsive formative assessment requires teachers to connect assessment expertise with the specific language needs of emergent bilingual learners. These findings reinforce the need to treat fairness as a contextual as well as a psychometric concern.

Culturally Responsive Assessment (CRA) provides one theoretical basis for addressing this challenge. CRA recognises that assessment design, administration, interpretation, and use should be responsive to students' cultural and linguistic experiences. Cultural validity is strengthened when assessment draws on learners' everyday experiences, allows appropriate agency, pays careful attention to language use, and avoids treating dominant cultural practices as neutral standards (Baidoo-Anu et al., 2024). A recent systematic review likewise identifies culturally relevant assessment as a means of improving the transparency and cultural validity of assessment for diverse learners (Ezeh & Kehinde, 2026).

Assessment for Learning (AfL) provides a complementary perspective because it positions assessment as an ongoing process for eliciting evidence, giving feedback, and adapting teaching rather than merely assigning scores. In K-12 EFL settings, formative assessment has been associated with improved teaching and learning when teachers possess adequate assessment literacy and when practices are adapted to contextual and cultural conditions (Zhang et al., 2024). Universal Design for Learning (UDL) further strengthens this perspective by encouraging multiple means of engagement, representation, and action or expression, thereby reducing barriers created by a single mode of demonstrating learning (CAST, 2024). In multilingual classrooms, these principles support flexible response formats, visual scaffolds, multimodal tasks, and differentiated pathways to demonstrate competence.

A further asset-based lens is provided by Funds of Knowledge theory, which views learners' household, community, linguistic, and cultural experiences as resources for learning rather than deficits (Moll et al., 1992). When assessment tasks connect with students' traditions, family experiences, and familiar cultural practices, they can provide more authentic opportunities to demonstrate knowledge. This position is consistent with culturally responsive teaching, which argues that learners' cultural knowledge and communicative practices should be incorporated into instructional and assessment decisions (Gay, 2010). Translanguaging extends the same logic to multilingual learners by allowing them to draw strategically on their full linguistic repertoires to comprehend tasks, formulate ideas, and demonstrate understanding (García & Wei, 2014).

Recent international scholarship demonstrates growing attention to these issues in primary education. In the United States, Yang (2025) documented linguistically responsive formative assessment with emergent bilingual pupils, while Rosheim et al. (2024) identified cultural and linguistic limitations in literacy assessment. In Australia, Dobinson et al. (2025) showed how monolingual classroom norms can render children's linguistic identities invisible. In Vietnam, Duong and Khuong (2025) found that culturally responsive assessment in ethnically diverse primary schools depended not only on teacher attitudes but also on institutional support. A 2026 scoping review of primary education concluded that linguistically responsive AfL remains a relatively small field and that pedagogical translanguaging, asset-based assessment, and teacher-researcher collaboration are promising directions (Robinson-Jones et al., 2026).

Indonesian primary education provides a particularly important setting for this discussion because linguistic diversity intersects with regional languages, Bahasa Indonesia, and English as a foreign language. Recent Indonesian research has similarly shown that valuing home languages and cultural identities can increase engagement and support more authentic demonstrations of language ability, while institutional pressures and standardised expectations may constrain teachers' implementation of culturally responsive assessment (Prastikawati et al., 2025). Nevertheless, context-specific evidence from urban multicultural primary schools remains limited.

Denpasar offers a distinctive setting for examining this problem. As the capital of Bali and a major urban and tourism centre, its schools bring together local Balinese communities, families from other Indonesian islands, children from mixed marriages, and expatriate families. Consequently, teachers may encounter students who differ in home language, cultural norms, prior educational experience, and familiarity with English. In such settings, the same assessment task can generate different levels of accessibility and meaning for different learners. This makes Denpasar an analytically important case for understanding how teachers negotiate the tension between common curriculum expectations and culturally and linguistically diverse learner profiles.

Previous studies have often examined either the limitations of standardised assessment, culturally responsive pedagogy, formative assessment, or multilingual scaffolding as separate issues. Less attention has been given to how primary-school teachers experience assessment challenges and construct practical responses within the same local cases, particularly in an Indonesian multicultural urban context. This study addresses that gap by examining both sides of the assessment process: the challenges teachers face and the adaptive strategies they employ. The study therefore asks: (1) What challenges do teachers encounter when conducting language assessment in multicultural primary classrooms in Denpasar? and (2) What strategies do teachers use to make language assessment more responsive and equitable?

Conceptual Framework

The conceptual framework in Figure 1 treats multicultural classroom diversity as the contextual condition that can generate linguistic, cultural, and resource-related assessment challenges. Teachers respond through differentiated and culturally responsive strategies, while institutional support and assessment literacy shape the extent to which these strategies can produce fairer, more valid, and more accessible opportunities for students to demonstrate learning.

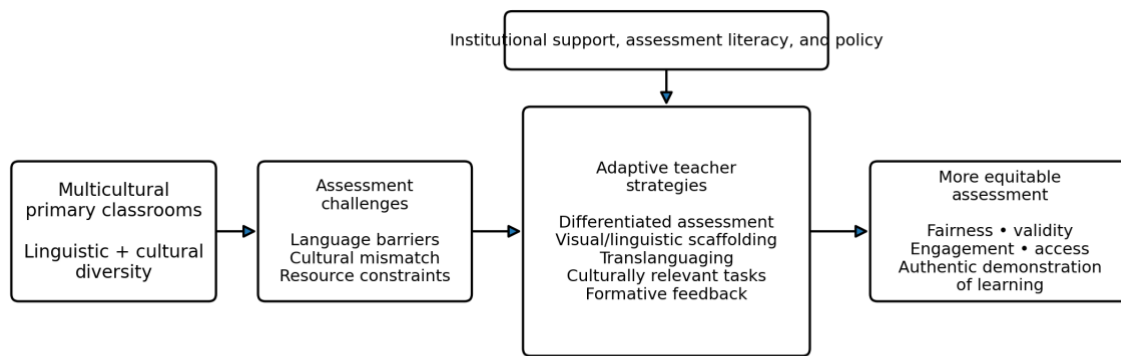


Figure 1. Conceptual framework linking classroom diversity, assessment challenges, responsive strategies, and equitable outcomes.

2. Method

This study employed a qualitative case study design because the research sought an in-depth, context-sensitive understanding of how teachers experience and respond to language-assessment challenges in multicultural primary classrooms. Qualitative case study is appropriate when a phenomenon is examined within its real educational setting and when participants' perspectives, classroom practices, and contextual conditions are central to interpretation (Creswell, 2014; Yin, 2018). The purpose was not statistical generalisation but analytical understanding of assessment practices in a specific multicultural urban context.

The study was conducted in public and private primary schools in Denpasar, Bali, selected because their classrooms included students from varied ethnic, linguistic, religious, and family backgrounds. Twelve teachers were purposively selected on the basis of their experience teaching in multicultural classrooms and their direct involvement in classroom assessment. This criterion ensured that participants had first-hand experience of the linguistic and cultural issues examined in the study.

Participant/context feature	Information reported
Number of participants	12 teachers
School sector	Public and private primary schools
Location	Denpasar, Bali
Sampling	Purposive sampling
Selection criterion	Experience teaching in multicultural classrooms and involvement in classroom assessment

Data were generated through three complementary sources. First, semi-structured interviews explored teachers' understandings of language-assessment principles, the linguistic and cultural difficulties they encountered, the adaptations they made, and the institutional support available to them. Second, classroom observations focused on how assessment was implemented and how students engaged with assessment tasks. Third, lesson plans, assessment rubrics, and student portfolios were examined as documentary evidence. These sources were used for triangulation so that interview accounts could be considered alongside observed practice and written assessment artefacts.

Ethical Considerations. Participant quotations are reported without personal identifiers, and the manuscript does not name individual teachers or participating schools.

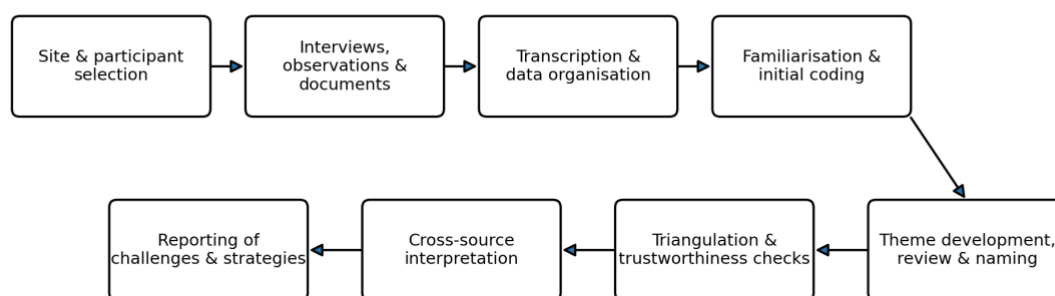


Figure 2. Research procedure and thematic-analysis workflow.

Data analysis followed Braun and Clarke's (2006) thematic-analysis framework. The researchers first familiarised themselves with the interview, observation, and document data; generated initial codes; grouped related codes into candidate themes; reviewed the themes against the data corpus; defined and named the themes; and produced the final analytic account. For example, statements concerning difficulty understanding English test instructions were coded as language-comprehension barriers and grouped within the broader theme of linguistic diversity and comprehension gaps, while references to flexible oral, written, visual, or project formats were grouped under differentiated assessment.

Trustworthiness was guided by Lincoln and Guba's (1985) criteria of credibility, transferability, dependability, and confirmability. Cross-source comparison and transparent thematic development were used to strengthen the credibility of interpretation.

3. Result

The findings draw on interview data, classroom observations, and document analysis and are organised around the two research questions: (1) challenges teachers encountered in language assessment and (2) strategies used to respond to those challenges. The analysis identified recurrent patterns rather than ranking themes numerically. Documentary data were used as corroborative evidence during theme review, while the quotations below represent the verbatim excerpts available in the source manuscript.

Teachers' Challenges in Language Assessment

Three closely related challenges emerged: linguistic and comprehension barriers, cultural misalignment in interpreting participation and performance, and limited culturally responsive assessment resources. Teachers described situations in which students appeared to understand lesson content but had difficulty accessing or responding to assessment tasks because the language of the task exceeded their English proficiency.

Excerpt 1 : 'Some students cannot comprehend the test instructions clearly. They come from homes where English is not their first language.'

The excerpt indicates that language proficiency can become an unintended barrier to demonstrating subject knowledge. Teachers therefore had to distinguish between difficulty with the construct being assessed and difficulty understanding the language used to

communicate the task. This challenge was especially salient for younger students and pupils whose home language differed from English.

Teachers also reported that culturally shaped participation behaviours could be interpreted inconsistently within common classroom rubrics. Quietness, indirectness, hesitation in public speaking, and avoidance of eye contact were mentioned as behaviours that could be read as low participation even when they reflected culturally appropriate forms of respect or restraint.

Excerpt 2 : 'When the assessment rubric sees not speaking up in class as low participation, but it is seen as being respectful in some ethnic groups.'

This mismatch created a risk that participation scores would reflect familiarity with dominant classroom interaction norms rather than language ability alone. Teachers therefore viewed cultural interpretation as part of assessment judgement, particularly in speaking, participation, and performance tasks.

Excerpt 3 : 'Most test books or assessment guidelines we use are generic. They do not consider students from non-mainstream backgrounds.'

A third challenge concerned the limited availability of assessment materials that represented the linguistic and cultural diversity of learners. Teachers described many test books and guidelines as generic and reported insufficient support for adapting them. The problem was therefore not only the presence of standardised materials but also the limited time, training, and institutional resources available for redesigning tasks in culturally responsive ways.

Teachers' Strategies in Addressing Challenges

Teachers responded to these challenges through four broad strategies: differentiated assessment, visual and linguistic scaffolding, culturally relevant task design, and more frequent formative assessment. These strategies were not treated as separate techniques; teachers combined them according to learners' age, language proficiency, task demands, and classroom context.

Excerpt 4 : 'I allow students to choose their assessment format. Some students present verbally, some of them write, and others draw. This type of assessment helps students show their strengths better.'

Differentiated assessment gave students more than one way to demonstrate learning. Teachers adjusted task complexity, allowed additional time when necessary, and used oral, written, visual, or project-based formats. Such flexibility enabled teachers to gather evidence of learning from students who might be disadvantaged by a single written format.

Excerpt 5 : 'I include visual instructions and use students' home language if needed, specifically for new or younger students.'

Visual and linguistic scaffolding was also used to make task expectations more accessible. Teachers used pictures, simplified instructions, gestures, and, when needed, students' home languages. Students could discuss or develop initial ideas using their first language before producing the final response in English. In this way, home-language support functioned as a bridge to comprehension rather than as a replacement for the target-language outcome.

Excerpt 6 : 'When students write or present about their traditions, they are more likely to do better and so is the quality of their language use.'

Teachers further connected assessment topics to students' traditions, family experiences, and cultural practices. They reported that familiar content increased students' willingness to participate and helped them produce richer language. Culturally relevant

tasks therefore served both an engagement function and an evidentiary function by creating conditions in which students could demonstrate what they knew through meaningful contexts.

Excerpt 7 : 'I give verbal feedback weekly, not only marks. It helps students and parents understand the progress without anxiety'

Finally, teachers increased the use of low-stakes assessment and regular verbal feedback. Rather than relying only on marks, they used feedback to communicate progress to students and parents and to reduce anxiety around assessment. Across the themes, the strategies point to an assessment approach that is flexible in mode, responsive to language needs, attentive to cultural meaning, and oriented towards supporting learning over time.

Table 1. Summary of themes, evidence, and interpretations

Area	Theme/sub-theme	Representative evidence	Interpretive meaning
Challenge	Linguistic/comprehension barriers	"Some students cannot comprehend the test instructions clearly..."	Task language may mask actual knowledge.
Challenge	Cultural misalignment	"Not speaking up in class" may be scored as low participation although it can signal respect.	Participation rubrics may privilege dominant interaction norms.
Challenge	Limited responsive resources	"Most test books or assessment guidelines we use are generic."	Teachers need adaptable materials and professional support.
Strategy	Differentiated assessment	Students may present, write, draw, or complete projects.	Multiple response modes widen opportunities to demonstrate learning.
Strategy	Visual/linguistic scaffolding	Visual instructions and home-language support are used when needed.	Scaffolding improves access to task meaning.
Strategy	Culturally relevant tasks	Students write or present about traditions and familiar experiences.	Cultural relevance can support engagement and richer language production.
Strategy	Formative feedback	Weekly verbal feedback is used in addition to marks.	Low-stakes feedback supports progress monitoring and reduces anxiety.

4. Discussion

The findings show that equitable language assessment in multicultural primary classrooms depends on whether assessment allows students to demonstrate the intended construct without unnecessary linguistic or cultural barriers. The difficulty some pupils experienced in understanding English instructions illustrates a threat to validity: when task language becomes more demanding than the ability being assessed, scores may partly reflect comprehension of instructions rather than the target skill.

This interpretation is consistent with Rosheim et al. (2024), who found that literacy assessments can disadvantage multilingual learners when linguistic and cultural relevance is weak, and with Yang (2025), who showed that responsive formative assessment requires teachers to connect assessment expertise with learners' specific language needs. Robinson-Jones et al. (2026) similarly conclude that linguistically responsive AfL in primary education should make space for learners' home languages and multimodal resources.

The cultural-misalignment theme extends this validity problem beyond language proficiency. Behaviours such as quietness, hesitation, or limited eye contact can carry different meanings across cultural communities; consequently, participation rubrics may become culturally loaded when they equate one interaction style with engagement or competence. Culturally Responsive Assessment therefore requires teachers to examine not only what is assessed but also the cultural assumptions embedded in criteria and scoring practices. Baidoo-Anu et al. (2024) argue that cultural validity involves attention to students' everyday experiences, agency, and language use. This perspective also aligns with Gay's (2010) culturally responsive teaching framework and with Funds of Knowledge theory (Moll et al., 1992), both of which position learners' cultural and community resources as assets for teaching and assessment rather than deviations from a dominant norm.

The reported shortage of culturally inclusive materials and professional support demonstrates that responsive assessment cannot be sustained through teacher creativity alone. Similar constraints have been identified internationally. Duong and Khuong (2025), studying ethnically diverse primary schools in Vietnam, found that support structures were a major condition for culturally responsive assessment, while Prastikawati et al. (2025) reported that institutional pressures and standardised expectations can limit responsive practices in Indonesian primary English classrooms. The present findings therefore broaden the issue from individual teacher competence to school-level assessment capacity. Access to adaptable materials, collaborative planning time, assessment-literacy development, and supportive policy are necessary if culturally responsive practices are to become routine rather than exceptional.

Differentiated assessment emerged as one practical response to diversity. Allowing students to present, write, draw, or complete performance and project tasks can reduce the bias introduced by a single response mode and can provide more complete evidence of learning. This practice is compatible with Universal Design for Learning, which recommends multiple means of action and expression and encourages educators to address barriers associated with particular modes of communication (CAST, 2024). It also strengthens Brown's (2004) principle of validity because the assessment format is less likely to obscure the intended competence. Differentiation should nevertheless preserve common learning goals and transparent criteria so that flexibility in expression does not become inconsistency in judgement.

Translanguaging and linguistic scaffolding were equally important because they enabled students to use familiar linguistic resources to access the meaning of assessment tasks. García and Wei (2014) conceptualise translanguaging as the strategic use of the learner's full linguistic repertoire rather than the simple alternation of separate languages. In the present study, home-language discussion, visual support, and simplified instructions functioned as bridges towards an English-language response. This reflects the asset-based orientation described in recent primary-assessment research (Robinson-Jones et al., 2026).

At the same time, translanguaging presents practical challenges: teachers may not share students' home languages, excessive support may alter the construct being assessed, and some learners may perceive different accommodations as unequal. Responsive use therefore requires clear decisions about when linguistic support improves access and when the target language itself is the construct to be measured.

The use of culturally relevant task content further indicates that assessment can become more authentic when students are invited to draw on traditions, family experiences, and familiar social practices. Such tasks activate funds of knowledge and reduce the distance between school assessment and learners' lived worlds (Moll et al., 1992). International work in Australia has shown the importance of making children's

linguistic and cultural identities visible rather than allowing monolingual norms to dominate classroom experience (Dobinson et al., 2025). A systematic review of culturally relevant assessment likewise reports positive evidence for culturally grounded practices in supporting diverse learners' assessment outcomes (Ezeh & Kehinde, 2026). In Denpasar, cultural relevance is particularly important because teachers work at the intersection of Balinese cultural traditions, inter-island mobility, and internationally influenced educational environments.

The teachers' use of regular verbal feedback and low-stakes assessment highlights the value of Assessment for Learning. Formative assessment is most useful when evidence of learning is used to adjust teaching and when feedback helps learners understand current performance and next steps. A recent systematic review of K–12 EFL education found that formative assessment can improve teaching and student outcomes, but its effectiveness depends on teacher assessment literacy and contextual conditions (Zhang et al., 2024). In multicultural classrooms, AfL is especially valuable because it gives teachers repeated opportunities to distinguish language-development needs from content-learning needs, monitor the effectiveness of scaffolds, and revise assessment tasks before high-stakes judgements are made.

School leadership and institutional policy shape whether these teacher strategies can be implemented consistently. Schools can support responsive assessment by providing shared banks of adaptable tasks, opportunities for moderation and collaborative rubric design, professional learning on multilingual assessment, and policies that permit reasonable flexibility in response mode and linguistic scaffolding. Without such structures, assessment equity depends too heavily on individual teachers and may vary considerably across classrooms.

Several practical implications follow. Teacher professional development should integrate assessment literacy with culturally responsive pedagogy, AfL, UDL, and translanguaging rather than treating them as separate topics. Curriculum developers should provide examples of culturally adaptable tasks and rubrics that distinguish the target construct from incidental language demands. School leaders should schedule collaborative assessment review and moderation, while policymakers should allow context-sensitive accommodations and encourage the development of locally relevant assessment resources. These measures would move culturally responsive assessment from an individual classroom strategy towards a sustainable school practice.

Limitations. The study is bounded by its qualitative case-study design, a sample of 12 teachers, and a limited number of public and private primary schools in one city. The findings therefore provide contextual and analytical insight rather than statistical generalisation. The study reports teacher perspectives but does not include student or parent interviews, and it does not provide quantitative frequency data for the themes. Future research could compare schools across Indonesian regions, incorporate student and parent perspectives, examine the effects of specific accommodations on assessment validity and learner outcomes, and investigate how school leadership influences the sustainability of responsive assessment practices.

Taken together, the findings suggest that equitable assessment emerges from the interaction of four elements: recognition of linguistic diversity, cultural responsiveness in task and rubric design, flexible and formative opportunities to demonstrate learning, and institutional support that enables teachers to sustain these practices. No single strategy is sufficient on its own. Differentiation without clear criteria may weaken comparability; translanguaging without construct clarity may alter what is being measured; and culturally relevant tasks without institutional support may remain isolated innovations. Equity is

therefore best understood as a coordinated assessment ecology in which technical quality and cultural responsiveness reinforce one another.

5. Conclusion

This study demonstrates that language assessment in multicultural primary classrooms in Denpasar is shaped by the interaction of linguistic diversity, culturally patterned participation, assessment-resource limitations, and institutional conditions. Teachers experienced difficulty when the language of assessment obscured students' actual knowledge, when common participation criteria conflicted with cultural norms, and when generic assessment materials offered limited support for diverse learners. In response, teachers used differentiated assessment formats, visual and linguistic scaffolding, translanguaging, culturally relevant tasks, and regular formative feedback. These strategies enabled students to access assessment tasks through more than one pathway and created more meaningful opportunities to demonstrate learning.

The study contributes to research on multilingual assessment by examining challenges and responsive strategies within the same multicultural primary-school cases in Denpasar. Its central implication is that fairness cannot be achieved by standardisation alone. Equitable assessment requires construct clarity, cultural validity, flexible modes of expression, and institutional support for teachers' assessment literacy. Schools should therefore provide professional development, collaborative assessment moderation, and adaptable assessment resources, while curriculum and policy should permit context-sensitive accommodations that preserve common learning goals. Future studies should include students' and parents' perspectives, compare different Indonesian regions, and examine how specific culturally and linguistically responsive practices influence assessment validity, learner engagement, and achievement over time.

6. References

- Baidoo-Anu, D., Liu, L., Cisterna, D., & Song, Y. (2024). Cultural validity: Promoting cultural responsiveness in classroom assessment. *Diaspora, Indigenous, and Minority Education*, 18(3), 160–181. <https://doi.org/10.1080/15595692.2023.2193883>
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests*. Oxford University Press.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Brown, H. D. (2004). *Language assessment: Principles and classroom practices*. Pearson Education.
- CAST. (2024). *Universal Design for Learning Guidelines version 3.0*. <https://udlguidelines.cast.org>
- Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). SAGE.
- Dobinson, T., Steele, C., & Winkler, G. (2025). From English-only to only English: Exploring children's linguistic and cultural identities in a primary school classroom. *Language and Education*, 39(5), 1107–1128. <https://doi.org/10.1080/09500782.2024.2433675>
- Duong, L. T., & Khuong, P. D. (2025). Culturally responsive assessment practices in ethnic minority semi-boarding primary schools in Vietnam: A mixed-methods study. *Frontiers in Education*, 10, 1670277. <https://doi.org/10.3389/feduc.2025.1670277>

- Ezeh, C. C., & Kehinde, O. J. (2026). Promising practices for culturally relevant assessment: A systematic review. *Diaspora, Indigenous, and Minority Education*, 20(1), 53–69. <https://doi.org/10.1080/15595692.2024.2341930>
- García, O., & Wei, L. (2014). *Translanguaging: Language, bilingualism and education*. Palgrave Macmillan.
- Gay, G. (2010). *Culturally responsive teaching: Theory, research, and practice* (2nd ed.). Teachers College Press.
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. SAGE.
- McNamara, T. (2001). Language assessment as social practice: Challenges for research. *Language Testing*, 18(4), 333–349. <https://doi.org/10.1177/026553220101800402>
- Moll, L. C., Amanti, C., Neff, D., & González, N. (1992). Funds of knowledge for teaching: Using a qualitative approach to connect homes and classrooms. *Theory Into Practice*, 31(2), 132–141. <https://doi.org/10.1080/00405849209543534>
- Prastikawati, E. F., Curle, S., Adeoye, M. A., & Vivekanantharasa, R. (2025). Culturally responsive assessment in Teaching English to Young Learners: Valuing home languages and identities. *Teaching English to Young Learners in Indonesia (TEYLIN)*.
- Robinson-Jones, C., Duarte, J., Chik, A., & Wang, H.-C. (2026). Linguistically responsive assessment for learning in primary education: A scoping review. *Journal of Multilingual and Multicultural Development*. Advance online publication. <https://doi.org/10.1080/01434632.2026.2613680>
- Rosheim, K. M., Tamte, K. G., & Froemming, M. J. (2024). Reducing inequalities inherent in literacy assessment of multilingual learners. *Reading Psychology*, 45(8), 735–751. <https://doi.org/10.1080/02702711.2024.2359922>
- Yang, X. (2025). Linguistically responsive formative assessment for emergent bilinguals: Exploration of an elementary teacher's practice in a math classroom. *International Multilingual Research Journal*, 19(1), 67–90. <https://doi.org/10.1080/19313152.2024.2339757>
- Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). SAGE.
- Zhang, H., Ge, S., & Mohd Saad, M. R. B. (2024). Formative assessment in K–12 English as a foreign language education: A systematic review. *Heliyon*, 10(10), e31367. <https://doi.org/10.1016/j.heliyon.2024.e31367>