



Castledown

 OPEN ACCESS

Language Education & Assessment

ISSN 2209-3591

<https://www.castledown.com/journals/lea/>

Language Education & Assessment, 9, 104539 (2026)

<https://doi.org/10.29140/lea.2026.104539>

EFL Teachers as Intercultural Gatekeepers: Designing Culturally Responsive Materials with GenAI in Indonesia



TOMMY HASTOMO^a 
FITRIA NUR HAMIDAH^b 
EVA NURCHURIFIANI^c 
KARTIKA S^d 
LYSTIANA NURHAYAT HAKIM^e 
SUKMA SEPTIAN NASUTION^f 
PIPIT MULIYAH^g 

^aSTKIP PGRI Bandar Lampung, Indonesia
Universitas Negeri Malang, Indonesia
tomhas182@gmail.com

^bPoliteknik Negeri Malang, Indonesia
Universitas Negeri Malang, Indonesia
fitria.nur@polinema.ac.id

^cSTKIP PGRI Bandar Lampung, Indonesia
churifiani@gmail.com

^dUIN Raden Intan Lampung, Indonesia
kartika@radenintan.ac.id

^eUniversitas Negeri Malang, Indonesia
Universitas Perjuangan Tasikmalaya, Indonesia
tianaghea@gmail.com

^fUniversity of Otago, New Zealand
Universitas Pamulang, Indonesia
sukmaseptian.nasution@postgrad.otago.ac.nz

^gInstitut Agama Islam K.H Sufyan Tsauri Majenang-Cilacap, Indonesia
pipitmuliyahoke@gmail.com

Abstract

Interculturality and culturally responsive pedagogy are gaining priority in Indonesian primary English as a Foreign Language (EFL) as learner diversity increases and Generative AI (GenAI) becomes part of everyday teaching. Yet little empirical work explains how EFL teachers use GenAI to design materials and tasks, or what language and ethics checks they perform. This qualitative case study examines teachers' perceptions and their strategies for maintaining linguistic appropriateness, cultural fit, and inclusion in GenAI-supported EFL resources. Twelve in-service primary English teachers were selected through purposive maximum-variation sampling across regions, school sectors, experience, and AI proficiency. Data included a GenAI-assisted design task with think-alouds, stimulated-recall

Copyright: © 2026 Tommy Hastomo, Fitria Nur Hamidah, Eva Nurchurifiani, Kartika S, Lystiana Nurhayat Hakim, Sukma Septian Nasution & Pipit Muliyah. This is an open access article distributed under the terms of the Creative Commons Attribution Non-Commercial 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within this paper.

interviews, and digital artifacts (prompts, outputs, and final materials). Reflexive Thematic Analysis was used to trace how teachers moved from initial prompts to revised materials and what checks they applied before considering the outputs suitable for pupils. Results show that GenAI speeds drafting and widens ideation, but teachers operate as intercultural gatekeepers: they contextualize prompts, refine outputs repeatedly, verify level and accuracy, evaluate cultural portrayals, and screen for bias to avoid generic or Western-leaning content while facing policy ambiguity, privacy risks, and uneven AI skills. The study recommends teacher training on culturally responsive prompting, critical checking of AI outputs, clearer school-level guidance on ethics and privacy, and better access to digital infrastructure across school contexts.

Keywords: Culturally responsive pedagogy, EFL materials, Generative AI, interculturality, TEFL

Introduction

Interculturality and cultural responsiveness have become foundational priorities in primary education as classrooms grow more plural through globalization and migration. Teachers are expected to develop intercultural competence that supports constructive dialogue, nurtures mutual understanding, and affirms the identities of learners from varied backgrounds (Hicham et al., 2025). At the same time, children encounter an expanding array of cultural narratives through global media, which makes it essential to equip them with the critical capacities to interpret multiple perspectives, interrogate power relations, and cultivate respect across difference (Dzerviniks et al., 2024). Well-designed intercultural pedagogy advances equity by disrupting discriminatory patterns and widening access to meaningful participation, while culturally responsive teaching has been associated with gains in motivation, social skills, and academic performance among ethnically diverse students (Siregar et al., 2023). This agenda aligns with broader commitments to social justice and invites schools to create environments where all students feel seen, supported, and able to succeed.

Artificial intelligence (AI) systems broadly refer to computer-based systems designed to simulate aspects of intelligent behavior through techniques such as machine learning, algorithmic processing, and natural language processing (Akgun & Greenhow, 2021). Within this broader category, GenAI refers to computational techniques capable of generating seemingly new and meaningful content, such as text, images, or audio, from patterns learned from training data (Feuerriegel et al., 2024). The rapid diffusion of GenAI intensifies both the promise and the responsibility of this work. On the one hand, GenAI can personalize learning by producing adapted texts, tasks, and immediate feedback, and it can increase efficiency in lesson design through tools that accelerate ideation and content development (Cui et al., 2025). On the other hand, it introduces risks that educators must manage, including biased or inaccurate outputs, privacy concerns related to data use, and challenges to academic integrity when assignments can be generated with minimal effort (Andewi et al., 2025). These conditions make it important to examine how teachers combine intercultural aims with the use of GenAI in everyday material design. As demographic diversity rises and AI capabilities expand, educators who combine critical intercultural aims with thoughtful adoption of GenAI are better positioned to design inclusive, equitable, and effective learning experiences that prepare children to thrive in an interconnected world (Santhini et al., 2025).

To understand how educators can exercise this responsibility in culturally diverse classrooms, it is necessary to clarify the intercultural principles that guide their use of GenAI. Intercultural competence refers to a multidimensional capacity to communicate and collaborate effectively across cultural difference, integrating cultural–linguistic knowledge, interactional skills, and affective orientations such as

empathy, openness, and reduced ethnocentrism (Schelfhout et al., 2022). Moreover, culturally responsive pedagogy mobilizes learners' cultural repertoires as instructional resources, aligning content and methods with students' sociocultural contexts to strengthen engagement, agency, and achievement (Meihami, 2023). To operationalize these aims, teachers draw on interculturally rich materials that foreground diverse texts and representations and invite perspective taking, comparison, and dialogue instead of privileging a single dominant viewpoint. GenAI can contribute by generating and adapting multimodal resources, proposing task ideas, and translating or contextualizing content for specific audiences, enhancing efficiency and broadening linguistic and cultural access (Mahmoudi-Dehaki & Nasr-Esfahani, 2025). Accordingly, this study examines these design principles in primary EFL classrooms, where English-learning materials mediate language development and intercultural understanding in everyday practice.

GenAI can expand teachers' design capacity by speeding up idea generation, offering varied approaches that spark creativity, and freeing more time for instruction (Moorhouse et al., 2024). It also supports differentiation by tailoring tasks to different abilities and learning preferences and by enabling culturally responsive materials that reflect students' backgrounds (Mahmoudi-Dehaki & Nasr-Esfahani, 2025). In addition, its language levelling features make resources more accessible through adaptation, translation, and localization. These advantages are tempered by amplified risks, since models may reproduce stereotypes or misrepresent cultures, generate inaccurate information, and obscure the provenance of content, all of which can weaken reliability and hinder the cultivation of critical inquiry (Darwin et al., 2024). In this landscape, teachers act as decisive gatekeepers who scrutinize outputs for accuracy, cultural sensitivity, and curricular relevance, contextualize materials to fit learners' cultural realities, ensure developmental suitability, and guide students to interrogate sources and claims (Riser, 2025). The promise of efficiency, inclusivity, and accessibility can be realized only when educators actively manage these hazards and model critical engagement so that classroom resources remain culturally safe, trustworthy, and pedagogically sound.

This study is framed by Indonesian curricular priorities that emphasize character education and multicultural values, while highlighting persistent gaps between policy aspirations and classroom enactment. Although national guidance promotes civic commitment and unity, everyday instruction can default to a narrow narrative that overlooks the country's ethnic and religious diversity (Pratiwi et al., 2023). The curriculum also endorses tolerance, intercultural communication, and recognition of cultural plurality, yet teachers often face limited institutional backing and restricted pedagogical flexibility when operationalizing these aims (Mariyono, 2024). Implementation is further complicated by highly diverse classrooms and by textbooks that underrepresent some minority communities, alongside regional resource disparities that produce uneven provision across schools (Waziana et al., 2025). Primary education is especially consequential because early exposure to difference shapes enduring dispositions toward inclusion and respect. In response, many primary teachers adapt or create materials to match local realities, particularly in under-resourced contexts, but sustained professional learning and targeted support remain vital for integrating multicultural and character education consistently (Ahamat & Kabilan, 2022).

In primary classrooms, intercultural and culturally responsive material design now intersects with GenAI, which can personalize content to students' cultural backgrounds and deepen engagement (Mahmoudi-Dehaki & Nasr-Esfahani, 2025). These benefits depend on teachers using professional judgment to embed cultural knowledge and pedagogical goals in the final materials. At the same time, implementation remains uneven because many educators lack sustained professional development and because cultural sensitivity is not consistently embedded in underlying models, leaving teachers to navigate resource constraints and design challenges (Gou, 2025). Teachers regard AI as a tool to increase efficiency and personalize learning, but they remain concerned about reliability, bias, and

overreliance on automated assessment (Bewersdorff et al., 2025). These concerns require robust ethical frameworks and sustained human oversight to ensure fairness, transparency, and the integrity of culturally responsive pedagogy.

Existing scholarship on GenAI in education largely concentrates on secondary and tertiary contexts (Bewersdorff et al., 2025; Darwin et al., 2024; Fountoulakis, 2024; Marzuki et al., 2023; Slamet, 2024; Wulyani et al., 2024), which leaves limited insight into how primary EFL teachers perceive and use these tools for intercultural and culturally responsive design and obscures their specific needs, constraints, and opportunities. In addition, few studies describe the actual intercultural communicative competence (ICC)-oriented content and EFL task types that teachers create in everyday practice, since many accounts emphasize conceptual possibilities or tool functions rather than classroom enactment. Procedures for assuring language accuracy, cultural appropriateness, and inclusivity in non-Western, multilingual contexts such as Indonesia are also insufficiently documented, including how teachers craft prompts, apply review criteria, and engage community members. These gaps call for studies that document teachers' actual prompts, AI-generated outputs, revisions, and checking practices in primary EFL classrooms. Positioning GenAI as a material-development partner for EFL, this study conceptualizes interculturality as ICC and examines teacher agency as teacher as intercultural and linguistic gatekeeper who evaluates language accuracy, cultural appropriateness, pedagogy fit, and ethics. This study is significant in three respects. Theoretically, it extends research on culturally responsive pedagogy and GenAI-supported language education by conceptualizing primary EFL teachers as intercultural and linguistic gatekeepers who actively evaluate cultural representation, language accessibility, pedagogical fit, and ethical risks in AI-generated materials. Practically, it shows how teachers use specific prompts, revise language level, check cultural representation, protect pupil data, and use student feedback before adapting GenAI-supported materials for classroom use. These strategies are also summarized in a one-page practice-oriented checklist to support classroom implementation. At the policy level, the study highlights the need for clearer institutional guidance, sustained professional development, and equitable access to digital infrastructure so that GenAI integration supports culturally inclusive primary education rather than widening existing disparities (Akgun & Greenhow, 2021; Mahmoudi-Dehaki & Nasr-Esfahani, 2025; Moorhouse et al., 2024). Accordingly, this study addresses the following research questions:

1. How do Indonesian primary EFL teachers perceive the use of GenAI to design culturally responsive and interculturally rich EFL materials?
2. What types of intercultural EFL tasks and EFL materials do teachers develop with GenAI, and how do they ensure linguistic accuracy, age/level appropriateness, cultural appropriateness, and inclusivity?

Literature Review

Interculturality and Culturally Responsive Pedagogy in Primary EFL

Interculturality in primary EFL education entails recognizing cultural plurality, sustaining dialogue across difference, and cultivating mutual respect through purposeful communication (Dusi et al., 2017). Within this orientation, culturally responsive pedagogy treats learners' cultural and linguistic histories as epistemic resources rather than background variables, strengthening instructional relevance and equity in language learning (Meihami, 2023). Teaching is designed to affirm identity, connect curriculum to local meanings and everyday experience, and invite critical engagement with multiple perspectives (Abacioglu et al., 2020). The aim extends beyond representational inclusion to encompass academic development, cultural competence, and reflective sociopolitical awareness that supports empathetic participation in diverse social contexts.

A growing body of evidence indicates that engagement and achievement improve when teachers build on community-based ways of knowing and situate EFL learning tasks within students' lived realities (Eguchi et al., 2021; Ramirez & Jimenez-Silva, 2015; Siregar et al., 2023). Strategies such as integrating indigenous knowledge, structuring perspective-taking activities, and using guided reflection can promote language growth, deepen conceptual understanding, and sustain participation in classroom discourse. These social-pedagogical moves also foster intercultural competence by enabling inclusive dialogue, constructive negotiation of difference, and heightened cultural self-awareness (Nameni, 2022). When embedded consistently, they contribute to classroom climates characterized by cooperation, tolerance, and equitable participation alongside academic rigor.

In Indonesia, multicultural and multilingual education features prominently in policy conversations, yet enactment in primary EFL classrooms remains uneven due to limited institutional support, rigid curricular pressures, resource disparities, and variability in teacher preparation (Munandar, 2024; Siregar et al., 2023; Sugianto et al., 2022). These constraints can reinforce the dominance of Bahasa Indonesia while sidelining local languages and knowledge systems. Strengthening practice requires context-sensitive guidance, more equitable resourcing, and sustained professional learning in culturally responsive teaching for linguistically diverse settings. Nonetheless, fine-grained empirical accounts of how Indonesian primary teachers design and implement interculturally responsive EFL materials and tasks, particularly when drawing on emerging GenAI tools, remain limited, motivating the present study.

GenAI in Intercultural Material Design

Research spanning education, art, and design suggests that GenAI can strengthen intercultural material design by enabling personalization, adaptive scaffolding, and culturally responsive content development (Grange et al., 2025; Jauhiainen & Garagorry Guerra, 2024; Mahmoudi-Dehaki & Nasr-Esfahani, 2025). In learning settings, GenAI may help produce resources that better reflect learners' diverse identities, connect tasks to locally situated knowledge, and reduce linguistic barriers through translation and contextual framing, thereby supporting inclusion and collaborative engagement (Jauhiainen & Garagorry Guerra, 2024). In higher education, studies report its use in creating culturally inclusive case studies and design templates intended to foreground relevance and curb bias, while allowing instructors to tune materials to students' prior experiences and community practices (Mahmoudi-Dehaki & Nasr-Esfahani, 2025; Zhang et al., 2025).

Parallel work in creative domains indicates that GenAI can assist practitioners in integrating heritage motifs into visual artefacts, installations, and apparel, combining efficiency with cultural fidelity when guided by domain expertise (Spennemann, 2024). Cross-media narrative projects that blend text, images, and interactivity have also been used to re-present classical literature and cultural resources in more accessible formats, reducing cognitive load and expanding participation. Across these applications, a consistent design principle is that humans set culturally meaningful constraints, use GenAI for iterative exploration across traditions, and refine outputs through critique–revision cycles (Grange et al., 2025).

However, responsible deployment depends on clear ethical and operational safeguards, including protection of cultural authenticity, bias mitigation, privacy preservation, and sustained human oversight to review and adapt outputs to local community contexts (Akgun & Greenhow, 2021; Al-Abdullatif, 2025; Massala, 2023). Emerging directions further combine GenAI with immersive technologies and analytics to deepen personalization and facilitate international collaboration (Jauhiainen & Garagorry Guerra, 2024). Although emerging studies have examined GenAI-supported personalization of school learning materials, culturally responsive EFL curriculum development for

pre-teen learners, and the professional GenAI competence of language teachers (Jauhiainen & Garagorry Guerra, 2024; Mahmoudi-Dehaki & Nasr-Esfahani, 2025; Moorhouse et al., 2024), insufficient empirical attention has been paid to how Indonesian primary EFL teachers translate intercultural and culturally responsive principles into actual material-design decisions. This gap is consequential because GenAI outputs may default to generic, culturally misaligned, or linguistically inappropriate content unless teachers actively contextualize prompts, evaluate language accessibility, and screen representations for bias. Moreover, limited evidence is available on the specific intercultural tasks and artefacts teachers produce or the safeguards they apply before introducing GenAI-supported materials into primary classrooms. Addressing this gap, the present study examines teachers' situated design practices and their roles as intercultural and linguistic gatekeepers in a non-Western, multilingual EFL context.

Methods

Research Design

This study is situated within a qualitative paradigm to generate a contextually grounded account of a contemporary educational phenomenon that is best understood through participants' interpretations and situated decision-making (Creswell & Poth, 2018). Because the research questions focus on how primary teachers make sense of GenAI, how they use it, and how they judge cultural appropriateness when designing intercultural tasks, the inquiry prioritizes meaning, reasoning, and process rather than quantification. Qualitative methods are therefore suited to capturing rich descriptions of experience and the evolving nature of action in authentic settings. An inductive analytic approach is employed to derive patterns and themes from teachers' narratives and design outputs, emphasizing the rationales and sequences underpinning pedagogical choices rather than counts or associations (Miles et al., 2014).

A qualitative case study design is used to examine a phenomenon embedded in real-world practice, where boundaries between the activity and its environment are permeable (Yin, 2018). The case is conceptualized as a bounded process: Indonesian in-service primary teachers' participation in an AI-assisted material-design task in which GenAI is used to generate, evaluate, and iteratively refine intercultural and culturally responsive EFL materials. While recruitment draws on maximum variation across regions, geography does not define the case; the unit of analysis is the shared design practice and task context. This framing foregrounds the interaction among teacher agency, the GenAI tool, the design activity, and the pedagogical intention of cultural responsiveness as they culminate in digital artifacts. Accordingly, this study is best understood as an embedded single instrumental case study: the 'case' is the shared GenAI-assisted EFL material-design practice, while individual teachers provide embedded perspectives within the same case.

Participants

This study used purposive maximum-variation sampling to recruit information-rich participants whose backgrounds and experiences were closely aligned with the research aims (Patton, 2015). The final sample consisted of 12 in-service Indonesian primary teachers of English who met three inclusion criteria: (1) they were currently teaching English at the primary level, (2) they designed their own English-learning materials, and (3) they had experience using at least one text- or image-based GenAI tool in their instructional preparation. To ensure participants had sufficient classroom experience to reflect meaningfully on pedagogical decision-making, all had at least two years of full-time teaching experience. Administrators and pre-service teachers were excluded because the study focused specifically on active classroom practitioners. Consistent with maximum-variation logic within a single-case design, participants were recruited from seven cities/regions: Lampung ($n = 2$), Surabaya ($n = 2$),

Malang ($n = 2$), Pontianak ($n = 2$), Bali ($n = 2$), Banjarmasin ($n = 1$), and Salatiga ($n = 1$). The sample also included teachers from both public ($n = 8$) and private schools ($n = 4$), with a gender distribution of seven females and five males. Participants ranged in age from 27 to 45 years ($M = 34.6$), while teaching experience ranged from 3 to 20 years ($M = 10.1$), indicating a mix of early-career and highly experienced teachers. GenAI proficiency was identified through a pre-recruitment screening survey using a three-level self-rating (Novice, Intermediate, Advanced) based on practical use descriptors; the sample included 3 novice, 6 intermediate, and 3 advanced users. Reported tool use was varied, with ChatGPT most frequently used (10/12), followed by Gemini (4/12) and Canva Magic Studio (2/12). Table 1 presents the participants' demographic and professional profiles in detail.

Table 1 *Participant demographic profile*

ID	Gender	Age	Region/ city	School type	Years of experience	GenAI proficiency	GenAI tools used
P01	F	27	Lampung	Public	3	Intermediate	ChatGPT, Canva Magic Studio
P02	M	34	Surabaya	Private	10	Advanced	ChatGPT, Gemini
P03	F	41	Malang	Public	17	Intermediate	ChatGPT
P04	F	29	Pontianak	Public	5	Novice	ChatGPT
P05	M	38	Bali	Private	12	Advanced	Gemini
P06	F	33	Banjarmasin	Public	8	Intermediate	Gemini
P07	M	45	Lampung	Public	20	Intermediate	ChatGPT, Canva Magic Studio
P08	F	30	Salatiga	Public	6	Novice	ChatGPT
P09	M	36	Surabaya	Private	11	Advanced	ChatGPT, Gemini
P10	F	28	Malang	Public	4	Intermediate	ChatGPT
P11	F	43	Bali	Public	18	Intermediate	ChatGPT
P12	M	31	Pontianak	Private	7	Novice	ChatGPT

Instruments

This study uses a multi-method qualitative strategy to develop a comprehensive account of the case through methodological triangulation (Patton, 2015). Three complementary data sources are integrated to capture both teachers' design processes and the materials they produce: a concurrent think-aloud protocol during an AI-assisted task, a post-task semi-structured interview, and a systematic analysis of the resulting digital artifacts. The think-aloud documents real-time reasoning, uncertainty, and iterative revisions; the artifact trail preserves how outputs evolve from initial generations to refined instructional products; and the interview situates observable decisions within teachers' beliefs, concerns, and professional judgments. These streams enable cross-source comparison to refine and nuance interpretations, while providing a layered view of how teachers plan, evaluate, and refine intercultural and culturally responsive materials with GenAI.

Data collection is organized around three instruments. First, participants complete an authentic EFL material-design task using a GenAI tool of their choice to create a short English-learning activity that is engaging, culturally responsive to their pupils, and oriented toward intercultural understanding.

During the task, prompts are limited to encouraging continuous verbalization to minimize post-hoc rationalization and retain an accurate record of thinking-in-action. Second, a semi-structured interview follows immediately to examine perceived affordances and constraints of GenAI for culturally responsive design and to support stimulated recall by linking specific moments from the think-aloud to explicit rationales. Third, an artifact corpus is compiled, including full AI interaction logs (prompts and outputs), intermediate versions, and final materials. Comparative artifact analysis traces how teachers select, adapt, and screen AI-generated content against pedagogical aims and cultural criteria, revealing concrete strategies used to secure appropriateness and inclusivity. The complete data-collection instruments are provided in Appendices A–C, including the pre-recruitment screening form, AI-assisted EFL material-design task and concurrent think-aloud protocol, and semi-structured stimulated-recall interview guide. Appendix D presents the artifact analysis dimensions used to support qualitative comparison across prompts, raw outputs, intermediate versions, and final materials.

The trustworthiness of the qualitative instruments was supported through procedures addressing credibility and dependability. Credibility was strengthened by aligning each instrument with the research questions, using a consistent task sequence across participants, and employing stimulated recall to connect interview responses with observable moments in the material-design process. Dependability was supported through standardized task instructions, neutral think-aloud prompts, systematic cross-source comparison of think-aloud transcripts, interviews, and design artifacts, and an auditable record of methodological and interpretive decisions. Peer debriefing with external colleagues was also used to review emerging codes, category boundaries, and thematic interpretations (Nowell et al., 2017). The mapping of the research questions to the data sources is presented in Table 2.

Table 2 Mapping research questions to data sources

Research question	Primary data source	Secondary data source(s)
RQ1. How do Indonesian primary EFL teachers perceive the use of GenAI to design culturally responsive and interculturally rich EFL materials?	Semi-structured interview (post-task)	Think-aloud protocols and design artifacts (AI-assisted material design outputs)
RQ2. What types of intercultural EFL tasks and EFL materials do teachers develop with GenAI, and how do they ensure linguistic accuracy, age/level appropriateness, cultural appropriateness, and inclusivity?	Think-aloud protocol (real-time design process)	Artifact analysis (prompts, raw outputs, final materials) Semi-structured interview (explicit explanation of safeguards and criteria)

Data Collection

Fieldwork and data handling were conducted systematically and in line with ethical research standards. Participants were recruited through professional networks, Indonesian teacher associations, and coordination with school principals. Interested teachers received an information sheet describing the study aims, participation requirements, expected time, and honorarium, then completed a brief screening survey to confirm eligibility. Eligible participants provided written informed consent that specified procedures, data types (audio, screen recordings, and digital artifacts), confidentiality safeguards (pseudonyms and anonymization), voluntary participation with the right to withdraw without penalty, and secure scholarly use of the data. Time was reserved for questions prior to consent.

Data were collected remotely via a secure video-conferencing platform to accommodate geographic dispersion. Each session included an orientation and a short practice activity to familiarize participants

with the think-aloud technique. Teachers then completed an AI-assisted design task while sharing their screens and verbalizing their reasoning; the researcher recorded audio and screen activity and used minimal prompts to sustain verbalization. A semi-structured interview followed immediately to elicit perceptions and rationales, supported by stimulated recall of salient moments. Data management procedures included encrypted, password-protected storage with restricted access, prompt anonymization by removing identifying details, verbatim transcription checked against recordings, and secure archiving of all files to maintain an auditable record and support credibility and dependability.

Data Analysis

This study applies Reflexive Thematic Analysis (RTA) as developed by Braun and Clarke to identify and interpret patterned meaning across the dataset (Braun & Clarke, 2021). RTA offers a systematic yet theoretically flexible approach suited to context-sensitive educational inquiry. A reflexive stance is central: themes are treated as analytic constructions generated through the researcher's interpretive engagement rather than as objective features awaiting discovery. Accordingly, the researcher serves as the primary analytic instrument, prioritizing explanatory themes over descriptive topic catalogues.

Analysis followed Braun and Clarke's (2021) six recursive phases across an integrated corpus of think-aloud transcripts, interviews, and artifact notes. The process involved immersion in the data, systematic coding of semantic and latent features, and iterative clustering of codes into candidate themes. During familiarization, the researcher read the transcripts repeatedly, reviewed the artifact trail, and recorded initial reflexive memos. Semantic and latent codes were then generated across the integrated corpus to capture both explicit statements and underlying patterns of meaning. Related codes were grouped into candidate themes and compared across the three data sources. Themes were then refined for coherence and fit with the full dataset, defined and named to clarify their conceptual focus and interrelations, and reported through analytic narratives supported by carefully selected extracts aligned with the research questions. Changes in codes, candidate themes, and interpretive decisions were documented through coding memos and a reflexive audit trail. Figure 1 summarizes this recursive process, while Appendix E provides examples of the coding trail from data extracts to initial codes, code clusters, and themes. The researcher served as the primary coder and analytic instrument. To strengthen the dependability of

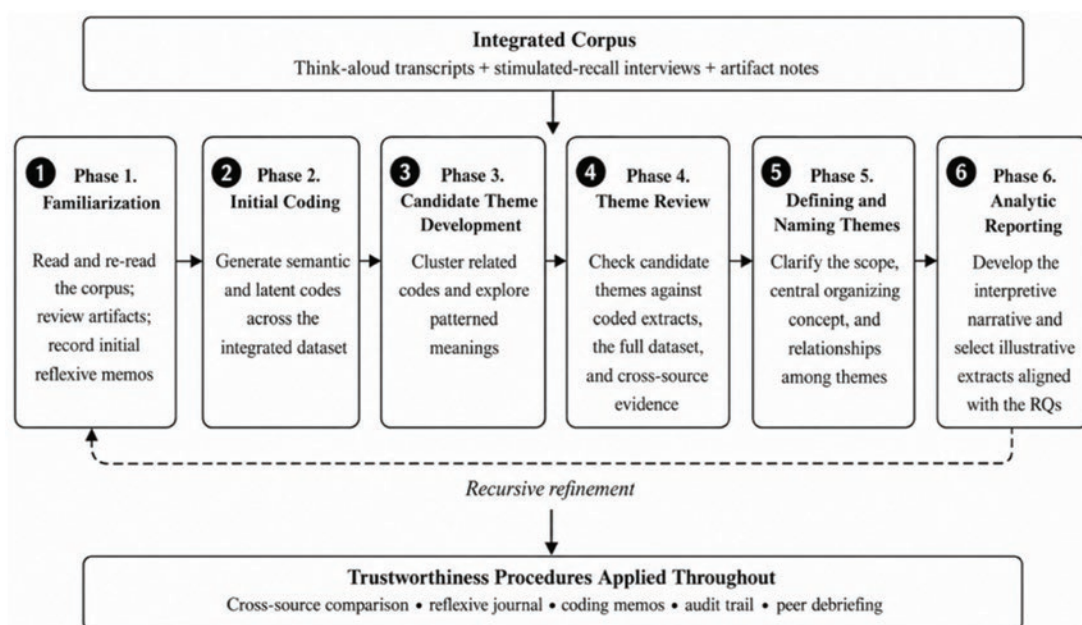


Figure 1 Recursive coding procedure in reflexive thematic analysis.

the analysis, emerging codes, category boundaries, and candidate themes were discussed with external peer debriefers. These discussions were used to question interpretations, refine theme boundaries, and check whether the analysis remained grounded in the think-aloud transcripts, interviews, and artifact notes. The study did not use a statistical inter-coder dependability index. Instead, dependability was supported through systematic cross-source comparison, reflexive memoing, an auditable record of coding decisions, and peer debriefing with external colleagues (Nowell et al., 2017).

Ethical Considerations

Ethical approval for the study was obtained from the Institutional Review Board (IRB) of one of the universities in Indonesia (Approval No. 13/LPPM/STKIP-PGRI/BL/Q/2026). All participants provided written informed consent prior to data collection. Confidentiality was protected through anonymization and secure data storage. To minimize participant burden, data collection was scheduled flexibly, breaks were offered when needed, and modest compensation was provided in accordance with institutional policy.

Results

This section reports findings in two subsections aligned with the research questions: teacher perceptions of GenAI (RQ1; Table 3) and the intercultural EFL tasks/materials teachers developed and safeguarded (RQ2; Table 4). GenAI adoption in primary education is neither uniform nor technologically deterministic; it unfolds as a sociotechnical process shaped by ongoing negotiations, tensions, and possibilities. Therefore, results across RQ1 and RQ2 converge into four overarching domains: shifts in teachers' roles and agency; teachers' perceptions of the engagement potential of GenAI-supported materials; opportunities and challenges for culturally responsive pedagogy; and ethical concerns involving bias, equity, and data privacy.

RQ1: Teacher Perceptions of GenAI

This section addresses the first research question. The findings reveal a continual negotiation between practical advantages and professional, ethical, and pedagogical reservations. Table 3 presents four themes: (1) duality of AI adoption with efficiency gains alongside concerns about professional identity, equity, and data privacy.

Table 3 *Thematic analysis summary of teacher perceptions of GenAI*

Theme	Sub-theme	Key findings
Duality of AI Adoption	Efficiency Gains vs. Professional Identity Concerns	Teachers use GenAI to ease workload in planning, drafting assignments, feedback, and communication, yet they report fear of displacement and a stigma of cheating that leads to cautious use.
The Teacher as Critical Mediator	Pedagogical Gatekeeping and Risk Management	Teachers act as negotiators and gatekeepers: they review and adapt outputs to maintain pedagogical control and attend to accuracy, reliability, and student data privacy.
Influence of Institutional Context	Policy Ambiguity and Teacher Anxiety	The presence or absence of clear and supportive policies shapes practice; unclear guidance is perceived as a barrier that creates anxiety and discourages innovation.
Pervasive Ethical Concerns	Algorithmic Bias and Educational Inequity	Teachers identify risks of gender, racial, cultural, and linguistic bias, and they are concerned GenAI may widen the digital divide in resource uneven settings.

Note. Themes were generated primarily from semi-structured interviews and were refined through cross-source comparison with think-aloud protocols and design artifacts from the AI-assisted material design task.

(2) teacher as critical mediator with gatekeeping and risk management through review and adaptation of outputs, (3) institutional context with policy ambiguity and anxiety that dampen innovation, and (4) pervasive ethical concerns with algorithmic bias and equity risks in settings where resources are uneven. To guide interpretation, the table provides a concise summary of themes, sub-themes, and key findings derived from think-aloud transcripts, semi-structured interviews, and design artifacts that capture how teachers perceive GenAI use.

Perceptions consistently balanced relief from workload with caution about role boundaries. Many teachers reported using GenAI for lesson planning, assignment drafting, feedback, and multilingual communication, while also expressing worry about displacement and the stigma of cheating that encouraged selective adoption. As one veteran teacher remarked, *“It is a double-edged sword. I can generate a week’s worth of reading comprehension questions in thirty minutes, which is incredible. But then I ask myself, am I still the one teaching, or is the machine?”* (P07) This illustrates the duality of AI adoption theme, where efficiency gains coexist with concerns about professional identity.

Across accounts, teachers framed themselves as active mediators rather than passive adopters. They described setting boundaries to preserve professional voice, emphasizing verification and adaptation so materials align with goals and learner needs. As one advanced user explained, *“I never take what Gemini gives me at face value. It is a starting point, a brainstorming partner. The final material has to have my pedagogical stamp on it, ensuring it is accurate and right for my kids.”* (P05) This reflects the teacher as critical mediator theme, particularly its gatekeeping dimension. This stance aligns with scholarship foregrounding teacher expertise, human oversight, and culturally responsive professional judgment in AI-supported language and K–12 education (Akgun & Greenhow, 2021; Meihami, 2023; Moorhouse et al., 2024; Riser, 2025).

Trust and risk management were central to this mediating work. Teachers reported persistent concerns about accuracy and reliability, noting that helpful suggestions can be accompanied by misleading phrasing that necessitates vigilant fact checking. They also described heightened attention to data privacy, from removing identifiable information to questioning institutional compliance. Unclear organizational guidance frequently produced anxiety and discouraged innovation, as a novice teacher remarked, *“We have not received any official guidelines. Everyone is a little scared to use it for graded assignments because we do not know what the official rule is. It feels like a grey area, so I mostly just use it for my own planning.”* (P08) This aligns with the institutional context theme, highlighting policy ambiguity and its chilling effect on innovation.

Ethical concerns extended to algorithmic bias, equity, and data protection. Teachers recognized that large models can reproduce gendered, racialized, cultural, and linguistic stereotypes, and they recounted outputs that defaulted to dominant narratives. *“I asked Gemini to create a story about a family celebration, and it gave me a very Western style birthday party. That does not reflect the experiences of most of my students here. You have to be so careful it does not erase their culture.”* (P06, Banjarmasin). This reflects the ethical concerns theme, with emphasis on algorithmic bias and cultural erasure. Equity issues such as costs, device access, and disparities between urban and rural areas were perceived as mechanisms by which GenAI may widen inequalities. Therefore, these findings suggest that teachers are discerning adopters who seek workload relief and pedagogical support while asserting agency, managing risk, and attending to ethics to sustain culturally responsive and context appropriate practice.

RQ2: Intercultural Content and Safeguards

This section addresses the second research question. The analysis indicates that teachers used GenAI to design intercultural artefacts that were differentiated not only by intercultural themes but also by

(a) skill focus (speaking, reading, writing, vocabulary), (b) task type (role-play, information gap, storytelling, guided reflection), and (c) language level (grade-based constraints and simplified proficiency cues) alongside explicit teacher checking and editing before classroom use. Table 4 summarizes four themes that connect these artefact features to teachers' safeguarding routines, triangulated across think-aloud transcripts, semi-structured interviews, and design artefacts.

Table 4 *Thematic analysis summary of intercultural content and safeguarding practices*

Theme	Sub-theme	Key findings
AI-Generated Content for Student Engagement	Skill focus & task types	Teachers use GenAI to draft speaking role-plays, short reading texts with questions, vocabulary practice, and writing frames that support low-stakes participation.
Leveraging GenAI for Culturally Responsive Teaching (CRT/ICC)	Localizing content vs. homogenization	Teachers localize settings, characters, and examples to invite perspective-taking, while countering generic or Western-centric defaults.
Ensuring Linguistic & Cultural Appropriateness	Prompt constraints + language/culture checks	Teachers constrain grade/age and task specifications in prompts, then audit vocabulary load, grammar, instructions, and cultural representations.
The Teacher's Role in the Loop	Editing, piloting, student feedback	Teachers iteratively revise outputs, trial small segments with learners, and refine prompts/materials before broader classroom use.

Note. Think-aloud protocols, semi-structured interviews, and design artifacts were used for triangulation.

Across skill areas, teachers often described speaking-oriented tasks (e.g., role-play dialogues) as salient outputs when working with GenAI. One teacher described a shy pupil who gained confidence after home rehearsal: *"I had a very shy student who never spoke. I used ChatGPT to create a role-play for home practice, and she volunteered next day."* (P01) Beyond dialogues, teachers used GenAI to draft information-gap exchanges and brief storytelling prompts that supported turn-taking, polite questioning, and respectful comparison. These tasks were framed as low-stakes practice for oral fluency.

Reading and vocabulary artefacts were typically short and tightly structured to fit EFL constraints. Teachers generated mini-texts paired with simple comprehension questions, culturally anchored word lists (foods, clothing, festivals), and picture-matching activities for quick recycling. For writing, they used GenAI to draft sentence frames, model paragraphs, and compare-and-contrast templates that supported identity-relevant expression while keeping instructions transparent. However, teachers stressed that intercultural value depended on deliberate localization rather than accepting generic outputs: *"The potential is huge. I can ask it to incorporate Balinese folklore into a science lesson. But if you are not specific, it just gives you generic and bland content."* (P11) This emphasis positioned task design as the driver of intercultural learning.

Language-level control functioned as a safeguarding layer enacted both upstream in prompts and downstream in edits. Upstream, teachers constrained outputs by stating grade/age (e.g., Grade 4; 8–10 years old), readability rules (short sentences, avoid idioms), and sometimes simplified proficiency cues (A1-like English). They also specified skill focus, task length (five to seven turns), and interaction format (pair work or home rehearsal) to avoid linguistically dense drafts. Prompt specificity was treated as essential for precision, as one teacher explained: *"You cannot just say a traditional Indonesian house. You have to specify a Joglo style house from Central Java with intricate wooden carvings, focusing on the pendopo area."* (P05) Downstream, teachers corrected grammar, simplified lexis, and rewrote directions into child-friendly language.

Cultural appropriateness and inclusivity were maintained through contextual prompting and teacher-in-the-loop review. Teachers inserted local place names, routines, and community norms, to better reflect classroom realities and to protect learners' identities. Several teachers described and explicitly requested non-stereotypical portrayals, then screened outputs for bias, overgeneralization, and tokenism. When content felt misaligned, they revised prompts or rewrote passages piloting a small excerpt with students before full use: *"After creating an AI-based activity, I test it with a few students, asking if it feels authentic. Their feedback is my key filter."* (P10) This iterative cycle positioned teachers as intercultural gatekeepers.

Discussion

This study examined how Indonesian primary EFL teachers perceive the use of GenAI for culturally responsive material design and how they develop, evaluate, and refine intercultural EFL tasks and resources. The findings show that GenAI supported teachers' material-design practices by accelerating idea generation and expanding the range of possible tasks. However, its educational value depended on teachers' capacity to contextualize prompts, assess language accessibility, evaluate cultural representations, and refine outputs before classroom use. Four interrelated findings are discussed: the coexistence of efficiency gains and professional concerns, the development of locally grounded EFL materials, the use of linguistic and cultural safeguards, and the influence of institutional and infrastructural conditions on teachers' adoption of GenAI.

First, teachers valued GenAI as a resource for reducing preparation time, generating ideas, and adapting learning activities. They used it to draft lesson materials, reading questions, speaking tasks, and feedback, while continuing to exercise professional judgment over the final outputs. This finding is consistent with Jauhiainen and Garagorry Guerra (2024), who highlighted the potential of GenAI to personalize school learning materials, and with Moorhouse et al. (2024), who emphasized the importance of teachers' professional GenAI competence. The present study extends these accounts by showing that efficiency gains were accompanied by concerns about pedagogical control and professional identity. Teachers viewed GenAI as a starting point for material development rather than as a substitute for their expertise. This suggests that the usefulness of GenAI in primary EFL education depends not only on its technical capacity to produce content but also on teachers' ability to evaluate whether that content is appropriate for their learners and classroom goals.

Second, teachers used GenAI to design varied EFL materials, including speaking role-plays, information-gap activities, short reading texts, vocabulary exercises, picture-matching tasks, sentence frames, and compare-and-contrast writing activities. These materials became more meaningful when teachers connected them to local settings, daily routines, cultural practices, foods, traditions, and community experiences. This finding supports culturally responsive pedagogy, which positions learners' cultural and linguistic repertoires as instructional resources rather than as peripheral background information (Meihami, 2023). It also aligns with Eguchi et al. (2021), who emphasized the importance of contextualizing AI-supported learning within culturally responsive approaches, and with Mahmoudi-Dehaki and Nasr-Esfahani (2025), who demonstrated the potential of GenAI to support culturally responsive EFL curriculum development for pre-teen learners. The present study adds a more detailed account of how primary EFL teachers transformed generic outputs into locally meaningful materials through contextualized prompting and iterative revision.

Third, teachers reported applying multiple safeguards before considering GenAI-supported materials suitable for classroom use. They specified learners' age or grade level, simplified vocabulary, reduced sentence length, avoided idiomatic expressions, checked grammar, clarified instructions, and adjusted task length and interaction formats. They also screened outputs for generic, stereotypical, tokenistic, or Western-centred cultural representations. These practices reinforce the argument that responsible AI integration requires sustained human oversight, particularly in K–12 education (Akgun &

Greenhow, 2021). They are also consistent with Al-Abdullatif's (2025) emphasis on awareness, ethics, evaluation, and responsible use as central dimensions of AI literacy. The findings add an EFL-specific perspective by showing that teachers' gatekeeping responsibilities operate across linguistic, pedagogical, intercultural, and ethical dimensions. Within the cases examined, insufficiently contextualized prompts sometimes produced generic or culturally inappropriate materials that required teacher revision.

Fourth, teachers' use of GenAI was shaped by institutional and infrastructural conditions. Participants expressed uncertainty about school-level guidance, learner privacy, the acceptable use of AI-generated materials, and unequal access to devices and internet connectivity. These concerns are consistent with Akgun and Greenhow's (2021) discussion of privacy and ethical challenges in K–12 AI use. They also reflect the broader Indonesian educational context, where the implementation of culturally responsive and technology-supported learning remains uneven because of differences in institutional support, teacher preparation, and resource availability (Ahamat & Kabilan, 2022; Mariyono, 2024; Sugianto et al., 2022). The findings indicate that responsible GenAI adoption cannot depend solely on individual teacher initiative. Clearer institutional guidance, sustained professional learning, and more equitable infrastructure are needed to support inclusive implementation.

This study contributes to the emerging literature by documenting primary EFL teachers' situated design decisions in a multilingual, non-Western context. Previous studies have largely focused on GenAI-supported personalization, teachers' professional GenAI competence, or culturally responsive curriculum development at a broader level (Jauhiainen & Garagorry Guerra, 2024; Mahmoudi-Dehaki & Nasr-Esfahani, 2025; Moorhouse et al., 2024). By integrating think-aloud protocols, stimulated-recall interviews, and artifact analysis, the present study shows how teachers moved from initial prompts and raw outputs to revised classroom materials. The findings conceptualize teachers as intercultural and linguistic gatekeepers who evaluate GenAI outputs through four interconnected lenses: linguistic accessibility, pedagogical suitability, cultural representation, and ethical responsibility. This perspective extends culturally responsive pedagogy by accounting for the evaluative work required when materials are generated or adapted through AI (Eguchi et al., 2021; Meihami, 2023).

The findings also have theoretical and pedagogical implications. Theoretically, they suggest that culturally responsive pedagogy in AI-supported learning environments cannot be reduced to the inclusion of cultural references in instructional materials. It also involves decisions about whose knowledge is represented, how cultural differences are framed, whether language demands are appropriate for young learners, and how potentially biased or generic outputs are revised (Eguchi et al., 2021; Meihami, 2023). Pedagogically, professional development should move beyond basic technical training and provide teachers with opportunities to practise context-sensitive prompting, language-level adjustment, bias screening, privacy protection, and iterative refinement (Akgun & Greenhow, 2021; Moorhouse et al., 2024). Institutions should also establish clear guidance regarding appropriate classroom uses of GenAI, data protection, and the review of AI-generated materials (Akgun & Greenhow, 2021).

To support classroom implementation, Supplementary Material A provides a one-page GenAI Intercultural Gatekeeping Checklist for Primary EFL Teachers. Derived from the teachers' reported practices, the checklist guides educators through five steps: contextualizing prompts, checking language accessibility, auditing cultural representation, protecting learner privacy, and piloting and refining GenAI-supported materials before classroom use. The checklist can be printed, shared digitally, or adapted for short professional-learning activities.

Conclusion

This study shows that the integration of GenAI in primary classrooms is a negotiated, sociotechnical process rather than a uniform technical shift. Teachers reported clear efficiency gains for planning,

adaptation, and feedback, yet they positioned themselves as critical mediators who review and refine AI outputs to protect accuracy, cultural relevance, and learner well-being. Teachers perceived that locally relevant, age-appropriate, and carefully scaffolded materials could support learner participation. However, this study did not directly examine pupils' engagement or learning outcomes. At the same time, teachers confronted risks related to bias, reliability, privacy, and uneven policy guidance, particularly in settings marked by infrastructural disparities. The analysis therefore reframes teacher agency as expanded rather than reduced, with educators acting as evaluators, ethical gatekeepers, and co-designers of culturally responsive materials that are locally attuned and pedagogically sound.

These findings have several implications. Practice should center sustained professional learning that develops critical AI literacy, ethical reasoning, culturally responsive prompt design, and routines for student feedback to validate cultural authenticity. As a practice-facing output, Supplementary Material A translates these recommendations into a printable five-step checklist for everyday GenAI-supported material design in primary EFL contexts. Policy should provide clear institutional guidance and invest in equitable infrastructure so that benefits do not amplify existing divides. This study also has several limitations that affect the transferability of the findings. It involved 12 Indonesian primary EFL teachers and a bounded set of remotely conducted think-aloud tasks, stimulated-recall interviews, and instructional artifacts. Participants' accounts may have been influenced by self-presentation, while the think-aloud procedure may have shaped their tool-use decisions. The material-design tasks did not provide direct evidence of pupils' learning outcomes or show how the materials would be enacted across different classroom settings. In addition, the rapid development of GenAI tools and institutional policies means that the reported practices reflect the technological and policy conditions available during data collection. Future research should examine classroom implementation, learner responses, longitudinal changes in teacher practice, and comparisons across educational levels and regional contexts.

Acknowledgments

The authors gratefully acknowledge financial support for this research and its article processing charge from several institutions, including the Indonesia Endowment Fund for Education (LPDP) Scholarship, Ministry of Finance of the Republic of Indonesia; the Indonesian Education Scholarship (BPI); the Centre for Higher Education Funding and Assessment (PPAPT), Ministry of Higher Education, Science, and Technology of the Republic of Indonesia; and the Indonesia Bangkit Scholarship (BIB) under the auspices of the Ministry of Religious Affairs of the Republic of Indonesia.

Declaration of AI and AI-assisted Technologies

While preparing this work, the authors used GenAI tools and Grammarly to support language editing and readability. The authors critically reviewed and revised all AI-assisted outputs and take full responsibility for the final content.

References

- Abacioglu, C. S., Volman, M. & Fischer, A. H. (2020). Teachers' multicultural attitudes and perspective taking abilities as factors in culturally responsive teaching. *British Journal of Educational Psychology*, 90(3), 736–752. <https://doi.org/10.1111/BJEP.12328>
- Ahamat, M. I. & Kabilan, M. K. (2022). Material adaptation among rural primary school English language teachers. *3L The Southeast Asian Journal of English Language Studies*, 28(1), 90–102. <https://doi.org/10.17576/3L-2022-2801-07>
- Akgun, S. & Greenhow, C. (2021). Artificial intelligence in education: Addressing ethical challenges in K-12 settings. *AI and Ethics*, 2(3), 431–440. <https://doi.org/10.1007/S43681-021-00096-7>

- Al-Abdullatif, A. M. (2025). Auditing AI Literacy competency in K–12 education: The role of awareness, ethics, evaluation, and use in human–machine cooperation. *Systems*, 13(6), 1–19. <https://doi.org/10.3390/systems13060490>
- Andewi, W., Waziana, W., Wibisono, D., Putra, K. A., Hastomo, T. & Oktarin, I. B. (2025). From prompting to proficiency: A mixed-methods analysis of prompting with ChatGPT versus lecturer interaction in an EFL classroom. *Journal of Studies in the English Language*, 20(2), 210–238. <https://so04.tci-thaijo.org/index.php/jsel/article/view/282318>
- Bewersdorff, A., Hornberger, M., Nerdel, C. & Schiff, D. S. (2025). AI advocates and cautious critics: How AI attitudes, AI interest, use of AI, and AI literacy build university students' AI self-efficacy. *Computers and Education: Artificial Intelligence*, 8, 1–12. <https://doi.org/10.1016/j.caeai.2024.100340>
- Braun, V. & Clarke, V. (2021). *Thematic analysis: A practical guide*. Sage.
- Creswell, J. W. & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches*. Sage.
- Cui, Y., Meng, Y., Qian, X. & Tang, L. (2025). Developing a college student feedback literacy scale for the GenAI context. *Assessment & Evaluation in Higher Education*, 1–14. <https://doi.org/10.1080/02602938.2025.2495048>
- Darwin, Rusdin, D., Mukminatien, N., Suryati, N., Laksmi, E. D. & Marzuki. (2024). Critical thinking in the AI era: An exploration of EFL students' perceptions, benefits, and limitations. *Cogent Education*, 11(1), 1–18. <https://doi.org/10.1080/2331186X.2023.2290342>
- Dusi, P., Roderigo, M. & Aristo, P. A. (2017). Teaching in our society: Primary teachers and intercultural competencies. *Procedia – Social and Behavioral Sciences*, 237, 96–102. <https://doi.org/10.1016/j.sbspro.2017.02.045>
- Dzerviniks, J., Ušča, S., Tarune, I. & Vindaca, O. (2024). Intercultural competence of teachers to work with newcomer children. *Education Sciences*, 14(8), 802. <https://doi.org/10.3390/educsci14080802>
- Eguchi, A., Okada, H. & Muto, Y. (2021). Contextualizing AI education for K-12 students to enhance their learning of AI literacy through culturally responsive approaches. *KI - Künstliche Intelligenz*, 35(2), 153–161. <https://doi.org/10.1007/s13218-021-00737-3>
- Feuerriegel, S., Hartmann, J., Janiesch, C. & Zschech, P. (2024). Generative AI. *Business & Information Systems Engineering*, 66(1), 111–126. <https://doi.org/10.1007/s12599-023-00834-7>
- Fountoulakis, M. St. (2024). Evaluating the impact of AI tools on language proficiency and intercultural communication in second language education. *International Journal of Second and Foreign Language Education*, 3(1), 12–26. <https://doi.org/10.33422/ijfsle.v3i1.768>
- Gou, D. (2025). The potential, challenges, and pathways of generative artificial intelligence in empowering the professional development of international Chinese language teachers. *Journal of Current Social Issues Studies*, 2(2), 104–115. <https://doi.org/10.71113/JCSIS.2025v2i2.104-115>
- Grange, C., Demazure, T., Ringeval, M., Bourdeau, S. & Martineau, C. (2025). The human-GenAI value loop in human-centered innovation: Beyond the magical narrative. *Information Systems Journal*. <https://doi.org/10.1111/isj.12602>
- Hicham, K., AlQbailat, N. M., Ismail, I. A., Qpilat, N. M., Al-Khawaldeh, N. N., Al-Shboul, O. K. & Masrar, F. Z. (2025). Interculturalizing ELT: Culture-based classes to enhance language skills and intercultural communicative competence dimensions. *Ampersand*, 14, 100221. <https://doi.org/10.1016/j.amper.2025.100221>
- Jauhiainen, J. S. & Garagorry Guerra, A. (2024). Generative AI and education: dynamic personalization of pupils' school learning material with ChatGPT. *Frontiers in Education*, 9, 1–12. <https://doi.org/10.3389/educ.2024.1288723>
- Mahmoudi-Dehaki, M. & Nasr-Esfahani, N. (2025). Utilising GenAI to create a culturally responsive EFL curriculum for pre-teen learners in the MENA region. *Education 3-13*, 53(7), 1175–1189. <https://doi.org/10.1080/03004279.2025.2502016>

- Mariyono, D. (2024). Indonesian mosaic: the essential need for multicultural education. *Quality Education for All*, 1(1), 301–325. <https://doi.org/10.1108/QEA-05-2024-0042>
- Marzuki, Widiati, U., Rusdin, D., Darwin & Indrawati, I. (2023). The impact of AI writing tools on the content and organization of students' writing: EFL teachers' perspective. *Cogent Education*, 10(2), 1–17. <https://doi.org/10.1080/2331186X.2023.2236469>
- Massala, K. (2023). Navigating bias and ensuring fairness equity unveiled in the AI-powered educational landscape. *Intelligence Artificielle et Technologie (Artificial Intelligence and Technology)*, 13(1), 37–41. <https://doi.org/10.7202/1107542ar>
- Meihami, H. (2023). Culturally responsive teaching in TEFL programs: Exploring EFL teachers' experiences. *English Teaching & Learning*, 47(1), 119–144. <https://doi.org/10.1007/s42321-022-00108-7>
- Miles, M. B., Huberman, A. M. & Saldaña, J. (2014). *Qualitative data analysis: A methods source-book*. Sage.
- Moorhouse, B. L., Wan, Y., Wu, C., Kohnke, L., Ho, T. Y. & Kwong, T. (2024). Developing language teachers' professional generative AI competence: An intervention study in an initial language teacher education course. *System*, 125, 103399. <https://doi.org/10.1016/j.system.2024.103399>
- Munandar, M. I. (2024). Interculturality and Islam in Indonesia's high-school EFL classrooms. *International Review of Applied Linguistics in Language Teaching*, 62(2), 301–323. <https://doi.org/10.1515/iral-2021-0200>
- Nameni, A. (2022). An investigation into Iranian employee's intercultural communicative competence: Does learning English as a foreign language help? *Journal of Intercultural Communication Research*, 51(4), 361–382. <https://doi.org/10.1080/17475759.2021.2006752>
- Nowell, L. S., Norris, J. M., White, D. E. & Moules, N. J. (2017). Thematic Analysis. *International Journal of Qualitative Methods*, 16(1). <https://doi.org/10.1177/1609406917733847>
- Patton, M. Q. (2015). *Qualitative research & evaluation methods*. Sage.
- Pratiwi, Y., Andajani, K., Suyitno, I., Ismail, A. & Prastio, B. (2023). Representing and implementing moral values to foreign students in Indonesian textbooks for learners other than Indonesians. *International Journal of Language Education*, 1(1), 58. <https://doi.org/10.26858/ijole.v1i1.36256>
- Ramirez, P. C. & Jimenez-Silva, M. (2015). The Intersectionality of culturally responsive teaching and performance poetry: Validating secondary Latino youth and their community. *Multicultural Perspectives*, 17(2), 87–92. <https://doi.org/10.1080/15210960.2015.1022448>
- Riser, P. (2025). Generative Artificial Intelligence (AI) in the high school English classroom and its effect on students' identity and teachers' practice. *English in Education*, 59(2), 176–192. <https://doi.org/10.1080/04250494.2025.2491361>
- Santhini, C., Ruba, M., Sakthi Priya, V. & Hemalatha, R. (2025). Intercultural communication and assessment practices in the integration of AI with education. In E. Meletiadou (Ed.), *Generative AI approaches to sustainable development in higher education* (pp. 253–280). IGI Global Scientific Publishing. <https://doi.org/10.4018/979-8-3693-5623-4.ch010>
- Schelfhout, S., Vandecasteele, R., De Maesschalck, S., D'hondt, F., Willems, S. & Derous, E. (2022). Intercultural competence predicts intercultural effectiveness: Test of an integrative framework. *International Journal of Environmental Research and Public Health*, 19(8), 4490. <https://doi.org/10.3390/ijerph19084490>
- Siregar, L. K., Mayuni, I. & Rahmawati, Y. (2023). Culturally responsive English teaching: Developing a model for primary school EFL teachers in Indonesia. *Educational Research*, 33(4), 1582–1600. <https://www.iier.org.au/iier33/siregar-abs.html>
- Slamet, J. (2024). Potential of ChatGPT as a digital language learning assistant: EFL teachers' and students' perceptions. *Discover Artificial Intelligence*, 4(1), 46. <https://doi.org/10.1007/s44163-024-00143-2>

- Spennemann, D. H. R. (2024). Will artificial intelligence affect how cultural heritage will be managed in the future? Responses generated by four GenAI models. *Heritage*, 7(3), 1453–1471. <https://doi.org/10.3390/heritage7030070>
- Sugianto, A., Kurniawan, E. & Sukyadi, D. (2022). Zooming in on the Indonesian EFL primary school students' intercultural sensitivity and their extroversion. *Studies in English Language and Education*, 9(1), 258–277. <https://doi.org/10.24815/SIELE.V9I1.21478>
- Waziana, W., Andewi, W., Wibisono, D., Hastomo, T. & Muslihudin, M. (2025). Exploring ChatGPT's impact on critical, creative, and reflective thinking skills: A mixed-methods study in an Indonesian EFL classroom. *Applied Research on English Language*, 14(4), 77–114. <https://doi.org/10.22108/are.2025.145896.2564>
- Wulyani, A. N., Widiati, U., Muniroh, S., Rachmadhany, C. D., Nurlaila, N., Hanifiyah, L. & Sharif, T. I. S. T. (2024). Patterns of utilizing AI-assisted tools among EFL students: Need surveys for assessment model development. *LLT Journal: A Journal on Language and Language Teaching*, 27(1), 157–173. <https://doi.org/10.24071/llt.v27i1.7966>
- Yin, R. K. (2018). *Case study research and applications: Design and methods*. Sage.
- Zhang, Q., Mohammad Ismail, M. I. R. & Zakaria, A. R. Bin. (2025). Enhancing intercultural competence in technical higher education through AI-driven frameworks. *Scientific Reports*, 15(1), 1–15. <https://doi.org/10.1038/s41598-025-03303-1>

Supplementary Material A

GenAI Intercultural Gatekeeping Checklist

For Primary EFL Teachers Designing Culturally Responsive Materials

Use this five-step checklist before using, sharing, or teaching materials generated with GenAI.

1

Contextualize the prompt

Add local places, foods, traditions, school routines, and learners' daily experiences.

Prompt cue: Create an A1-level dialogue for Grade 4 pupils in Indonesia using a local market setting.

2

Check the language level

Review vocabulary, grammar, sentence length, instructions, and task load for age and proficiency.

Ask: Can my pupils understand and complete this without unnecessary difficulty?

3

Audit cultural representation

Look for stereotypes, tokenism, Western-centered examples, or inaccurate cultural portrayals.

Ask: Does this material respect and accurately represent local and diverse cultures?

4

Protect learners

Do not enter students' names, photos, personal stories, addresses, or identifiable classroom data.

Rule: Use anonymized or fictional examples when prompting GenAI.

5

Pilot and refine

Test a small part of the material, collect student responses, then revise prompts and final materials.

Ask: What confused students, what felt authentic, and what needs simplification?

Key message: GenAI can support EFL material design, but teachers remain the final cultural, linguistic, and ethical gatekeepers.

Based on findings from Indonesian primary EFL teachers' GenAI-assisted material-design practices.

Appendix A. Pre-Recruitment Screening Form

Purpose

This brief form was used to confirm participant eligibility and obtain background information relevant to purposive maximum-variation sampling.

Section A. Professional Background

1. Are you currently teaching English at the primary-school level?
☐ Yes ☐ No
2. How many years of full-time teaching experience do you have?
_____ years
3. What is your age?
_____ years old
4. What is your gender?

5. What type of school do you currently teach in?
☐ Public school ☐ Private school
6. In which city or region do you currently teach?

7. Do you design or adapt your own English-learning materials for classroom use?
☐ Yes ☐ No

Section B. Experience with GenAI Tools

8. Have you used at least one text- or image-based Generative Artificial Intelligence (GenAI) tool when preparing English-learning materials?
☐ Yes ☐ No
9. Which GenAI tools have you used for instructional preparation? Select all that apply.
☐ ChatGPT ☐ Gemini ☐ Canva Magic Studio ☐ Other: _____
10. Briefly describe one example of how you have used GenAI in your teaching preparation.

Section C. Self-Rated GenAI Proficiency

11. Select the description that best represents your current proficiency in using GenAI tools for instructional preparation.
- ☐ Novice: I can enter basic prompts and use simple outputs, but I still need support to refine prompts or evaluate output quality.
 - ☐ Intermediate: I can revise prompts, adapt outputs for my pupils, and identify some issues related to accuracy, language level, or cultural appropriateness.
 - ☐ Advanced: I can iteratively refine prompts, compare outputs, evaluate linguistic and cultural appropriateness, and apply ethical safeguards when using GenAI-generated materials.

Appendix B. AI-Assisted EFL Material-Design Task and Concurrent Think-Aloud Protocol

Purpose

This task was used to examine how primary EFL teachers generated, evaluated, and refined culturally responsive and interculturally oriented learning materials with GenAI.

Material-Design Task

Using a GenAI tool of your choice, create a short English-learning activity for your primary-school pupils. The activity should be engaging, culturally responsive to your pupils, and oriented toward intercultural understanding. You may revise your prompts and the generated materials as needed. Please retain your prompts, generated outputs, intermediate versions, and final material.

Concurrent Think-Aloud Instructions

While completing the task, verbalize what you are thinking. Explain what you notice and what you are considering when you retain, revise, or remove particular elements. There are no correct or incorrect answers. The purpose is to understand your reasoning during the material-design process.

Minimal Researcher Prompts

When necessary, the researcher used neutral reminders to sustain verbalization:

- “Please continue thinking aloud.”
- “Could you explain what you are considering at this point?”
- “Please describe what you are noticing.”

The researcher did not recommend content or evaluate the quality of the generated materials during the task.

Materials Retained for Analysis

- full GenAI interaction logs;
- initial prompts;
- revised prompts;
- raw GenAI outputs;
- intermediate versions;
- final English-learning materials;
- screen recordings; and
- think-aloud audio recordings.

Appendix C. Semi-Structured Stimulated-Recall Interview Guide

Purpose

This post-task interview examined teachers' perceptions of GenAI-supported material design and explored the rationales underlying decisions observed during the concurrent think-aloud task. Follow-up questions were adapted to each participant's interaction log and final material.

1. How do you perceive the use of GenAI for preparing primary EFL materials?
2. What considerations guided your decisions when creating this activity?
3. How did you decide whether the generated language was appropriate for your pupils?
4. How did you decide whether the content was relevant to your pupils' cultural context?
5. Did you revise any part of the generated output? Please explain your reasons.
6. What checks do you normally perform before using GenAI-generated materials in class?
7. What challenges or concerns do you associate with the use of GenAI in primary EFL teaching?
8. What forms of institutional or professional support would help teachers use GenAI more responsibly?

Stimulated-Recall Prompts

The following prompts were used when referring to specific moments in the screen recording or interaction log:

- “At this point, you revised your prompt. Could you explain your reason?”
- “What influenced your decision to retain, change, or remove this element?”
- “What made this version more suitable for your pupils?”
- “Would you revise this material further before using it in class? Why?”

Appendix D. Artifact Analysis Dimensions

Purpose

These dimensions supported qualitative comparison of prompts, raw GenAI outputs, intermediate versions, and final materials. They were used as an analytic guide.

Section A. Artifact identification

Field	Description
Participant ID	Pseudonym assigned to the teacher
GenAI tool	ChatGPT, Gemini, Canva Magic Studio, or other tool
Grade or age group	Intended primary-school level
Skill focus	Speaking, reading, writing, vocabulary, or integrated skills
Task type	Role-play, information gap, storytelling, guided reflection, reading comprehension, picture matching, writing frame, or other type
Version	Initial prompt, revised prompt, raw output, intermediate version, or final material

Section B. Artifact Analysis Dimensions

Category	Guiding Questions	Notes and Evidence
Prompt development	How did the prompt change across iterations? Were context, age, language, or task specifications added or revised?	
Linguistic accessibility	Are the vocabulary, grammar, sentence length, and instructions appropriate for the intended pupils?	
Pedagogical fit	Does the material align with the intended learning objective and classroom context?	
Cultural contextualization	Does the material incorporate locally meaningful settings, practices, examples, or references?	
Intercultural orientation	Does the activity encourage dialogue, comparison, perspective-taking, or recognition of diversity?	
Cultural representation	Are any portrayals generic, inaccurate, stereotypical, tokenistic, or excessively centred on a dominant cultural perspective?	
Inclusivity	Are learners and cultural practices represented respectfully?	
Privacy	Does the material avoid identifiable pupil data?	
Teacher revisions	Which elements were retained, revised, added, or removed between versions?	
Cross-source memo	Which explanations from the think-aloud protocol or interview clarify the observed revisions?	

Section C. Comparative Analytic Memo

For each participant, record a brief analytic memo addressing the following questions:

1. How did the teacher's prompt change across iterations?
2. Which linguistic, pedagogical, cultural, or ethical concerns shaped the revisions?
3. What was retained, adapted, or removed from the raw GenAI output?
4. How did the final material differ from the initial output?
5. What patterns of teacher decision-making were visible across the artifact trail?
6. What additional practices, if any, were reported in the think-aloud protocol or interview but could not be directly observed in the artifacts?

Appendix E. Sample Coding Trail

The table presents illustrative examples of how selected data extracts were coded, grouped into code clusters, and developed into themes.

Table 1 *Sample coding trail*

Illustrative Data Extract	Initial Code	Code Cluster	Theme
"It is a double edged sword. I can generate a week's worth of reading comprehension questions in thirty minutes, which is incredible. But then I ask myself, am I still the one teaching, or is the machine?" (P07)	Efficiency gains accompanied by concern about teacher displacement	Efficiency gains versus professional identity concerns	Duality of AI Adoption
"I never take what Gemini gives me at face value. It is a starting point, a brainstorming partner. The final material has to have my pedagogical stamp on it, ensuring it is accurate and right for my kids." (P05)	AI output treated as a starting point requiring teacher verification and adaptation	Pedagogical gatekeeping and risk management	The Teacher as Critical Mediator
"I asked Gemini to create a story about a family celebration, and it gave me a very Western style birthday party. That does not reflect the experiences of most of my students here. You have to be so careful it does not erase their culture." (P06)	Generic Western-centred output creates a risk of cultural erasure	Algorithmic bias and cultural misrepresentation	Pervasive Ethical Concerns
"The potential is huge. I can ask it to incorporate Balinese folklore into a science lesson. But if you are not specific, it just gives you generic and bland content." (P11)	Local specificity is required to prevent generic outputs	Localizing content versus homogenization	Leveraging GenAI for Culturally Responsive Teaching
"After creating an AI-based activity, I test it with a few students, asking if it feels authentic. Their feedback is my key filter." (P10)	Student feedback used as an authenticity filter	Editing, piloting, and student-informed refinement	The Teacher's Role in the Loop