

TEACHER-AI DIALECTIC IN CULTURALLY SITUATED ENGLISH ASSESSMENT: TRANSFORMING MANDAR ETHNOLITERACY THROUGH HUMAN-AI CO-CREATION

Nurqadriyanti Hasanuddin¹, Hustiana², Amrang³, Ikhsan⁴

¹²³⁴ Universitas Sulawesi Barat, Indonesia

Email: hustiana@unsulbar.ac.id

ABSTRACT

English language evaluation in West Sulawesi has long relied on culturally foreign materials, producing a digital cultural gap that alienates students from Mandar identity, while an emerging AI-centric paradigm risks eroding teachers' pedagogical agency. This study examines the dialectic and negotiation of meaning undertaken by five English teachers in the Mandar region as they transformed local ethnoliteracy including the sandeq boat, the Metawe' politeness tradition, and the Boyang traditional house into interactive English evaluation instruments through a Human-AI Co-Creation strategy. Adopting a qualitative, document-based Participatory Design, this article reports findings from reflexive thematic analysis of five prompt-engineering logs (coded T1–T5) produced during a Human-AI Co-Creation Workshop involving four generative AI platforms (DeepSeek, ChatGPT, Claude, and Gemini). Analysis identified five interrelated themes: contextualizing prompts through curriculum and local culture, negotiating completeness of the evaluation structure, negotiating linguistic-level appropriateness, technical dialectic over AI system limitations, and validating substance to filter cultural hallucination. These themes were synthesized into a preliminary five-stage model of teacher–AI dialectic: Contextual Initiation, Initial AI Generation, Critical Teacher Audit, Negotiation and Iterative Revision, and Resolution. The findings indicate that teachers function as active co-designers rather than passive operators, and that the cultural accuracy of AI output depends heavily on the richness of the contextual input teachers supply at the outset. This document-analysis phase forms part of a broader study; teachers' phenomenological experience of the dialectic process will be examined separately through hermeneutic-phenomenological interview data in a subsequent publication.

Keywords: Ethnoliteracy; Hallucination filtering; Human-AI co-creation; Mandar culture; Negotiation of meaning; Prompt engineering

Introduction

English instruction at the local level frequently suffers from material alienation, in which the texts and evaluation instruments used draw predominantly on foreign cultural references. When teaching and evaluation materials disregard students' local context, learners struggle to construct meaning from the content presented (Yusniawati & Lestari, 2021). In West Sulawesi, the richness of Mandar

ethnoliteracy the seafaring tradition of the sandeq boat (Nianty et al., 2024), the Sayyang Pattudduq tradition, the Lipaq Saqbe textile heritage, and the Siri' code of honor has not been optimally utilized as a source of authentic, contextual English evaluation material. This underutilization is not confined to English instruction alone; a needs analysis of Mandar-language teaching in West Sulawesi high schools similarly found that although most schools had nominally adopted

regional-language curricula, only a small fraction of students were actually engaged in formal Mandarin instruction, with many acquiring the language informally through the home environment rather than through school-based channels (Muliastuti et al., 2025). These findings reinforce the broader claim that embedding local culture into language teaching, including English language teaching, has been shown to encourage students to interpret texts critically, strengthen cultural identity, and increase engagement with the learning process (Alakrash & Abdul Razak, 2021; Fitria, 2025).

Beyond language pedagogy, the depth of Mandarin local wisdom available as pedagogical raw material is also evident in literary studies of Mandarin culture, which identify recurring, teachable forms such as the architecture of the panggung house, communal gotong-royong practices surrounding house-building and relocation, and distinctive local foodways (Trang & Hieu, 2026). Such culturally embedded content, including the practice of collective cooperation and food traditions tied to specific communal rituals, illustrates the density of authentic material available for contextualizing evaluation instruments—material that, to date, remains largely untapped in English assessment design (Trang & Hieu, 2026). This condition constitutes a digital cultural gap in which English evaluation feels foreign because it does not draw on local context.

A preliminary study conducted by the research team found that teachers commonly copy evaluation items directly from textbooks and worksheets, reflecting a pedagogical and professional condition that falls short of producing interactive, locally grounded evaluation (Amrang & Hustiana, 2025; Hustiana et al., 2025). At the same time, the rapid adoption of generative AI in education has triggered a paradigm shift from human-centric to AI-centric practice, in which technology risks replacing rather than augmenting teachers' natural capacities (Raeburn, 2025). Left unchecked, this tendency risks standardizing all local cultures into a single, homogenized global flavor. The notion of Human-AI Co-Creation adopted in this study reflects a critical awareness that AI needs human partners in this case, English teachers acting as custodians of culture so that the resulting transformation does not become

"hallucinated" or stripped of the spirit of tradition (Ji et al., 2023; Liu, 2025). This is consistent with the view that collaboration between human creativity and machine intelligence is indispensable for preserving accuracy, integrity, and humanistic values in education in the digital era (Hadi et al., 2025).

A growing body of scholarship has begun to examine how teachers exercise agency when working with generative AI. Studies grounded in ecological and sociocultural models of agency show that language teachers actively negotiate, resist, and redirect AI output rather than adopt it uncritically (Jia & Pun, 2025; Lee et al., 2025), and that agency in AI-mediated language teaching is increasingly distributed across teachers, learners, and systems rather than located in any single actor (Godwin-Jones, 2024; Zhai, 2024). Parallel to this, conceptual frameworks of Human-AI Co-Creation describe iterative cycles of generation, evaluation, and revision between human and machine partners (Fragiadakis et al., 2024; Liu, 2025), while research on culturally responsive AI has begun to document how AI systems can be adapted to sociocultural contexts, though this literature remains largely conceptual or review-based and rarely draws on classroom-adjacent transcript data (Kaplan-Rakowski et al., 2023; Nguyen, 2025). Participatory design scholarship, meanwhile, has long argued that positioning teachers as design partners, rather than end users produces classroom technologies that better fit local practice (Musaeus et al., 2024; Nicholson et al., 2022), a principle with clear relevance to culturally situated assessment design, and awareness of AI hallucination as a risk to educational content has grown substantially across the field (Kasneci et al., 2023).

Despite this expanding literature, three gaps remain only partially addressed. First, teacher-agency studies on generative AI have concentrated on lesson planning and general instructional use, with far less attention to high-stakes evaluation design (Jia & Pun, 2025; Lee et al., 2025). Second, Human-AI Co-Creation frameworks have rarely been tested against culturally situated, minority-language contexts in which the risk is not merely factual hallucination but the erosion of a specific cultural identity (Ladson-Billings, 1995). Third, and most directly relevant to this article, almost none of this

literature examines the moment-to-moment dialectic, the turn-by-turn argument, correction, and negotiation, through which teachers actually produce a finished evaluation artifact with AI; most studies instead rely on retrospective interviews or surveys about AI use rather than documented interaction logs. Whereas prior studies have largely positioned teachers as passive operators of digital tools or examined their perceptions after the fact, this study positions teachers as active subjects observed in situ as they negotiate meaning with AI, retaining full authority to filter AI output so that cultural distortion is avoided. The novelty of this study lies in its methodological focus on the documented process of dialectic itself, the interactional process through which teachers “argue” with an algorithm while transforming Mandar ethno-literacy into evaluation instruments, a dimension seldom explored empirically in AI-in-education research generally, and not previously documented in the West Sulawesi context. This article specifically addresses the first research question guiding the broader study: how do English teachers enact dialectic and negotiate meaning within Human-AI Co-Creation as they transform Mandar ethno-literacy values into English evaluation instruments? Findings concerning teachers' phenomenological experience of this process (the study's second research question) will be reported in a subsequent publication once in-depth interview data reach saturation.

Method

Research Design

This study employs a qualitative approach with a Participatory Design that positions teachers as active co-designers, rather than mere end users, of the resulting evaluation instruments. The research data were obtained from direct observation of interaction logs (prompt-engineering logs) between teachers and various AI models, namely DeepSeek, ChatGPT, Claude, and Gemini. Specifically, this study analyzes five interaction logs (coded T1–T5) generated during a Human-AI Co-Creation Workshop, treating each log as a bounded case of the dialectic process as it unfolded in situ. This document-based analytical unit was deliberately chosen because prompt-engineering logs capture

the interaction turn-by-turn as it occurred, allowing the dialectic process to be examined directly rather than reconstructed retrospectively.

Participants and Ethical Considerations

Data collection involved five purposively sampled English teachers from the Mandar region, selected based on a minimum of two years of teaching experience in a Mandar cultural context and prior experience interacting with generative AI platforms in an instructional setting. This sample size is consistent with recommendations for qualitative case-based research and the principle of information power, whereby a small but highly relevant sample can achieve data saturation (Malterud et al., 2016). All participants provided informed consent prior to the workshop, were briefed on the voluntary nature of participation and their right to withdraw at any point, and were assigned the anonymized codes T1–T5 used throughout this article; no personally identifying information is reported. The broader study received internal ethical clearance through LPPM Universitas Sulawesi Barat prior to data collection, and workshop transcripts and derived data are stored on password-protected institutional storage accessible only to the research team.

Human-AI Co-Creation Workshop Procedure

The broader study proceeds through four stages: (1) Ethnographic Content Mapping to identify and validate Mandar ethno-literacy values suitable for classroom use; (2) a Human-AI Co-Creation Workshop in which teachers engage in prompt engineering and negotiate meaning with AI output, producing the documented interaction logs analyzed here; and (3) reflexive thematic analysis and synthesis across stages. In Stage 2, each teacher independently selected one validated ethno-literacy theme from Stage 1 and one of four generative AI platforms with which they already had prior familiarity (DeepSeek, ChatGPT, Claude, or Gemini), accessed through each platform's standard free-tier web interface. Sessions were largely unstructured beyond a brief task card asking teachers to produce a complete, classroom-ready English evaluation instrument

(items and a scoring rubric) aligned to their assigned grade level; researchers observed the session but did not intervene in prompt formulation. Session length ranged from approximately 45 to 90 minutes, and the number of documented interaction turns per case ranged from 6 to approximately 20. Because sessions used each platform's standard consumer interface rather than a version-pinned API, exact model version numbers were not systematically logged at the time of data collection; this is acknowledged as a limitation in the Conclusion, and version logging has since been added to the protocol for subsequent workshop cycles.

Data Analysis

Data were analyzed following the six-phase reflexive thematic analysis procedure (Clarke & Braun, 2017): familiarization through repeated reading of each transcript; initial open

coding of every teacher and AI turn, conducted independently by two members of the research team; searching for candidate themes across related codes; reviewing and refining themes against the full dataset, including reconciling coding disagreements through discussion rather than a fixed statistical agreement coefficient given the small case-based sample; defining and naming themes; and producing the thematic report integrated with the research question. The resulting codebook and a theme-by-artifact matrix (Table 3) are reported below to make the coding process transparent and to support cross-case comparison, as recommended for qualitative case analysis.

Findings and Discussion

The five artifacts analyzed vary in the AI platform used and the ethnoliteracy theme addressed, as summarized in Table 1

Table 1.
Profile of the Human-AI Co-Creation Artifacts

Code	AI Platform	Ethnoliteracy Theme	Output Produced
T1	DeepSeek	Lopi Sandeq (traditional Mandar boat) descriptive text “My Sandeq Boat”	Multiple-choice & fill-in items, scoring rubric, manual PDF workaround
T2	ChatGPT	Metawe’ the Mandar tradition of politeness and respect	Full evaluation set: MC, short answer, essay, multi-component rubric, discussion prompts, self-assessment
T3	Claude	Lopi Sandeq parallel descriptive text (document-based)	20 items across four sections, separate answer key, rubric, Word file
T4	DeepSeek	Boyang, Mandar traditional house and its ‘Dua Tassisara, Tallu Tammallaesang’ philosophy	Test blueprint, 10 MC + 5 essay items, answer key, rubric, learning objectives
T5	Gemini	Lopi Sandeq, A1-level children’s song built on boat-part vocabulary	Song lyrics, vocabulary glossary, interactive quiz, attempted LaTeX-based PDF instrument

Reflexive thematic analysis of the five artifacts produced five interrelated themes, discussed below.

Contextualizing Prompts through Curriculum and Local Culture

This theme emerged consistently across all five cases and marks the starting point of the dialectic: teachers did not enter the interaction as passive users but as designers who had already constructed a pedagogical frame. In T1 and T3,

for instance, teachers explicitly positioned the AI as a “professional/EFL English teacher for Grade VII junior high school in Indonesia” and specified detailed learning objectives, including target vocabulary related to the lopi sandeq. A similar pattern appeared in T4, where the teacher requested items “aligned with the learning objectives in the file,” together with a test blueprint. These actions show that the negotiation of cultural meaning effectively began before the AI produced a single sentence, at the moment the

prompt itself was formulated. The teacher functions as an initial curator, filtering which Mandar terms and philosophies are worth feeding into the AI (Jia & Pun, 2025).

Negotiating Completeness of the Evaluation Structure

This theme was most pronounced in T2 and T5, in which teachers acted as quality auditors of AI output. In T2, after ChatGPT produced an evaluation set that included a rubric only for the essay section, the teacher firmly objected: “why is the rubric only for the essay section.” This objection prompted the AI to overhaul the rubric to cover all evaluation components (multiple choice, short answer, essay, discussion, and self-assessment), aligning with the authentic, holistic assessment principles of the Merdeka Curriculum across knowledge, skills, and attitude domains. An escalating pattern also appeared in T5: the teacher initially accepted a simple interactive quiz from Gemini but subsequently raised the demand to a formal, single-file instrument comprising 10 multiple-choice items, 5 matching items, and 5 essay items with a complete rubric. Such negotiation confirms that structural completeness is not automatically produced by AI but must be actively “won” by teachers through repeated dialectic cycles.

Negotiating Linguistic-Level and Learner Appropriateness

This theme was explicitly identified in T5, which shows a layered negotiation concerning appropriateness for CEFR level A1 and junior high school learner characteristics. The teacher requested two rounds of adjustment first, to convert generic song lyrics into vocabulary-focused lyrics (“change the lyrics into something like learning vocabulary about sandeq”), and second, to ensure level appropriateness (“make me a song suitable for A1 level junior high school”) even after the AI had already claimed its output met that level. This repetition suggests that teachers do not simply trust an AI’s self-reported claims of appropriateness but demand re-verification a reflective form of dialectic directed at the machine’s epistemic authority. This is

consistent with recent evidence that LLMs’ self-reported confidence is often poorly calibrated and remains overconfident even when their output is inaccurate (Cash et al., 2026).

Technical Dialectic: Detecting and Resolving AI System Limitations

This theme reveals a dimension of dialectic seldom discussed in the conceptual Human-AI Co-Creation literature: negotiation over technical system limitations rather than cultural content. In T1, DeepSeek candidly acknowledged its inability to generate a PDF file directly and offered an alternative HTML-based workaround, requiring the teacher to take over part of the technical work outside the AI system. In T3, this process occurred internally on the AI side (Claude self-debugged a failed table-generation error in a Word document), a capability consistent with documented self-debugging behavior in large language models more broadly (Chen et al., 2023), which, although it did not involve verbal negotiation from the teacher, remained part of the broader dialectic of artifact production. Most notable is T5, where Gemini failed to compile LaTeX code into a PDF and explicitly stated, “I apologize, there seems to be a problem with the document code I created earlier, so it failed to compile correctly.” The T5 transcript ends at this point without a recorded resolution, indicating that not every Human-AI Co-Creation dialectic concludes smoothly within a single session; some instead leave friction that demands further patience and strategy from the teacher beyond the documented session.

Validating Substance and Filtering Cultural Hallucination

This theme relates directly to the study’s core concern: the teacher’s role in safeguarding cultural integrity against the risk of AI “hallucination” (Ji et al., 2023). Findings indicate that this filtering occurred predominantly in an implicit, preventive manner rather than a reactive, corrective one: teachers largely prevented cultural errors through highly specific, source-document-based prompts (T1, T3, T4) rather than correcting cultural mistakes after they had

already appeared. Illustratively, in T3, the distractor option generated for a question about what propels the lopi sandeq such as “engine/motor” was in fact consistent with the cultural fact that sandeq is propelled by sail and wind rather than an engine, showing that the AI can produce culturally accurate content when supplied with adequate source-document context by the teacher. In T4, the AI successfully formulated a higher-order essay question on the philosophy of “Dua Tassisara, Tallu Tammallaesang” precisely because that philosophy was already fully available in the teaching material the teacher had prepared and validated beforehand (an output of the Ethnographic Content Mapping stage). These findings reinforce the argument that the cultural quality and accuracy of AI output depend heavily on the quality of contextualization teachers provide at the outset, such that hallucination filtering effectively begins well before the AI produces an answer at the moment the teacher decides which information is worth feeding to the system.

Cross-Case Coding Evidence

To make the analytic process transparent and support cross-case comparison, Table 3 summarizes which of the five themes were coded in each artifact. The pattern shows that

contextualizing prompts (Theme 1) was the only theme present across all five cases, consistent with its status as the entry point of the dialectic; the remaining themes surfaced unevenly depending on the specific friction each teacher encountered with their chosen platform and content.

Table 2.
Theme–Artifact Coding Matrix

Theme	T1	T2	T3	T4	T5
1. Contextualizing Prompts	✓	✓	✓	✓	✓
2. Negotiating Completeness	–	✓	–	–	✓
3. Negotiating Linguistic-Level Appropriateness	–	–	–	–	✓
4. Technical Dialectic	✓	–	✓	–	✓
5. Validating Substance / Filtering Hallucination	✓	–	✓	✓	–

Toward a Preliminary Model of Dialectic Stages

Synthesizing the five themes above, this study proposes a preliminary five-stage model of teachers’ dialectic process in Human-AI Co-Creation, presented in Table 2. This model is provisional and will be refined through triangulation with hermeneutic phenomenological interview data in a subsequent phase of the study, particularly to surface the cognitive and emotional dimensions of teachers at each stage.

Table 3.
Preliminary Model of the Teacher–AI Dialectic Process

Stage	Description
1. Contextual Initiation	The teacher formulates a prompt that assigns the AI a pedagogical role, states learning objectives and basic competencies, and embeds specific Mandar cultural terms (e.g., lopi sandeq, Metawe’, Boyang, Tallu Tammallaesang).
2. Initial AI Generation	The AI produces a draft evaluation instrument, often with incomplete coverage or a format that does not yet fit classroom needs.
3. Critical Teacher Audit	The teacher re-reads the output critically, checking completeness across assessment domains, linguistic-level appropriateness, cultural accuracy, and document format.
4. Negotiation and Iterative Revision	The teacher issues follow-up instructions or corrections (e.g., “why is the rubric only for the essay section”; “change the lyrics into vocabulary-focused ones”), compelling the AI to explain its limitations and revise its output.
5. Resolution	The cycle ends in one of two ways: (a) a final product the teacher accepts as an authentic representation of Mandar values, or (b) an unresolved technical obstacle (e.g., a failed document compilation) that requires teacher intervention beyond the AI system.

This five-stage model confirms that Human-AI Co-Creation in this context is not a one-way process in which AI generates and teachers passively accept, but an iterative cycle that continues until the teacher not the AI declares the product finished and culturally and pedagogically valid. In four of the five cases (T1, T2, T3, T4), the cycle recorded closure at Stage 5(a). In one case (T5), the cycle was recorded as ending at Stage 5(b), an important finding indicating that teacher agency in safeguarding evaluation quality does not always correlate directly with the technical capacity of the AI system used.

These findings reinforce the view that teachers' role in collaborating with generative AI is shifting from a mere tool operator to an active co-designer with full agency over the direction and quality of output (Hadi et al., 2025; Liu, 2025). This shift aligns with the notion of human-centered AI, which asserts that technology should augment, not replace, natural human capacities (Raeburn, 2025). In this study's context, the capacity being augmented is not merely general pedagogical competence but, more specifically, teachers' capacity as custodians of Mandarin ethnoliteracy identity. The situation over the past five years (2021–2026) shows a trajectory relevant to these findings. In the early part of this period, scholarly attention centered on the urgency of integrating local culture into English teaching materials to counter the dominance of foreign content (Yusniawati & Lestari, 2021) and on teachers' digital literacy as a precondition for meaningful technology use (Alakrash & Abdul Razak, 2021). By the middle of the decade, scholarly focus shifted toward repositioning teachers as co-designers in AI-assisted task and assessment design rather than passive recipients of system output. Concurrently, awareness of AI hallucination risk the tendency of language models to produce plausible-sounding yet false or fabricated information (Ji et al., 2023) has grown substantially and become a serious concern across education and scholarship broadly, as reflected in recent guidance emphasizing layered verification of AI-generated content before it is used in instructional or scholarly contexts.

Positioned against this literature, the contribution of the five-stage model proposed here is threefold. First, existing Human-AI Co-Creation frameworks describe iterative cycles of generation, evaluation, and revision in largely

generic terms (Fragiadakis et al., 2024); the present model specifies a distinct Critical Teacher Audit stage (Stage 3) oriented explicitly toward cultural fidelity rather than only technical or structural quality, giving empirical texture to what culturally responsive AI scholarship has so far mostly discussed at the conceptual or systematic-review level (Nguyen, 2025). Second, ecological and ecological-agency accounts of teachers working with generative AI (Jia & Pun, 2025; Lee et al., 2025) identify iterational, projective, and practical-evaluative dimensions of agency but do not map these dimensions onto the concrete, sequential behaviors observable in artifact production; the present model grounds those abstract dimensions in an observable production sequence — from prompt formulation through resolution — that can be traced directly in interaction logs. Third, whereas participatory-design scholarship has argued for positioning teachers as co-designers of educational technology in general (Musaeus et al., 2024; Nicholson et al., 2022) and culturally relevant pedagogy scholarship has long argued that teachers must actively mediate curriculum content for cultural fidelity (Ladson-Billings, 1995), this study connects both arguments to a specific, high-stakes artifact, the English evaluation instrument and to a specific mechanism, hallucination filtering, through which that mediation occurs in practice.

The technical and cultural-dialectic themes identified here (Themes 4 and 5) contribute empirically to this discourse: the risks Mandarin English teachers face in practice involve not only cultural-content hallucination but also outright technical failure (such as the failed document compilation in T5), which demands a different form of vigilance and resilience. Furthermore, the pattern in Theme 5 that AI cultural accuracy depends heavily on the quality of source documentation teachers supply during contextualization corroborates prior findings that successful integration of local culture in English teaching hinges on teachers' ability to articulate local knowledge explicitly and systematically (Fitria, 2025), including knowledge of the *sandeq* boat as a Mandarin cultural heritage now threatened by modernization (Nianty et al., 2024). In other words, the richer and better validated the ethnographic input teachers provide, the smaller the space for AI to “hallucinate” culturally a

finding consistent with the research team's continuing work on local-culture-based, technology-supported English learning practices among Mandar youth (Amrang & Hustiana, 2025; Hustiana et al., 2025)(Amrang & Hustiana, 2025; Hustiana et al., 2025).

Conclusion

This study demonstrates that the dialectic and negotiation of meaning enacted by Mandar English teachers within Human-AI Co-Creation unfolds as a five-stage cycle centered on teacher agency rather than on the sophistication of the AI algorithm alone. Teachers initiate the dialectic from the moment a prompt is formulated, critically audit the completeness and accuracy of AI output, negotiate revisions repeatedly, and, in some cases, must accept that the negotiation does not necessarily culminate in full technical resolution within a single session. These findings position teachers not as passive consumers of AI-generated content but as active gatekeepers of Mandar cultural identity, whose contextualization choices at the outset substantially determine the cultural fidelity of the resulting evaluation instruments.

These findings should be interpreted in light of several limitations. First, the evidence base comprises only five purposively selected cases from a single cultural region and workshop, and the resulting model is offered as a preliminary, context-specific account rather than a generalizable claim about Human-AI Co-Creation. Second, the analysis draws exclusively on prompt-engineering logs; it does not yet incorporate the hermeneutic-phenomenological interview data that will speak to teachers' cognitive and emotional experience of the dialectic, and readers should not infer phenomenological findings from this article beyond what the documented artifacts directly support. Third, because sessions used each AI platform's standard consumer interface without systematic version logging, the findings describe teacher-AI dialectic under free-tier conditions available in September 2025 and may not generalize to other model versions, paid tiers, or subsequent system updates. Fourth, the study captures a single workshop session per teacher and does not address how these instruments performed once implemented in an actual classroom, nor the longer-term sustainability of the practice. Future research should triangulate

these findings with the forthcoming interview data, extend the sample across additional cultural themes and regions, and follow the resulting instruments into classroom use to examine their pedagogical validity beyond the design stage. Subsequent analysis of hermeneutic phenomenological interview data will extend these findings by illuminating the cognitive and emotional dimensions underlying this dialectic process.

References

- Alakrash, H. M., & Abdul Razak, N.** (2021). Technology-Based Language Learning: Investigation of Digital Technology and Digital Literacy. *Sustainability*, 13(21), 12304. <https://doi.org/10.3390/su132112304>
- Amrang, A., & Hustiana, H.** (2025). Exploration of English Learning Practices with Edtech and Local Culture Among Mandar Youth. *Journal of English Teaching, Applied Linguistics and Literatures (JETALL)*, 8(3), 184. <https://doi.org/10.20527/jetall.v8i3.23771>
- Cash, T. N., Oppenheimer, D. M., Christie, S., & Devgan, M.** (2026). Quantifying uncertainty: Testing the accuracy of LLMs' confidence judgments. *Memory & Cognition*, 54(2), 375–400. <https://doi.org/10.3758/s13421-025-01755-4>
- Chen, X., Lin, M., Schärli, N., & Zhou, D.** (2023). *Teaching Large Language Models to Self-Debug* (Version 2). arXiv. <https://doi.org/10.48550/ARXIV.2304.05128>
- Clarke, V., & Braun, V.** (2017). Thematic analysis. *The Journal of Positive Psychology*, 12(3), 297–298. <https://doi.org/10.1080/17439760.2016.1262613>
- Fitria, T. N.** (2025). Culture-Based Teaching: Integrating Indonesian Culture in English Language Teaching (ELT). *International Journal of Sustainable English Language, Education, and Science*, 2(2), 62–76. <https://doi.org/https://doi.org/10.71131/xbqdn589>
- Fragiadakis, G., Diou, C., Kousiouris, G., & Nikolaidou, M.** (2024). *Evaluating Human-AI Collaboration: A Review and Methodological Framework* (Version 2). arXiv. <https://doi.org/10.48550/ARXIV.2407.19098>
- Godwin-Jones, R.** (2024). Distributed agency in second language learning and teaching through generative AI. *Language Learning & Technology*, 28(2), 5–31. <https://doi.org/10.64152/10125/73570>
- Hadi, J. K., Latifah, H., Fuadi, D. A., Fauzan, F., Christiana, Y., Hidayat, T., & Rifa'i, R.**

- (2025). *Kolaborasi Manusia-Mesin Dalam Pendidikan: Strategi Guru Beradaptasi Dengan Teknologi AI*. 6329–6333. <https://doi.org/10.31004/riggs.v4i2.1583>
- Hustiana, H., Nugroho, D. A., Asdar, A., Hasanuddin, N., Junaedi, A., & Fajriani, F.** (2025). Prospective Teachers' Perceptions Of English Area Training: Enhancing 21st-Century Competencies. *Journal of English Teaching, Applied Linguistics and Literatures (JETALL)*, 8(3), 129. <https://doi.org/10.20527/jetall.v8i3.22284>
- Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y. J., Madotto, A., & Fung, P.** (2023). Survey of Hallucination in Natural Language Generation. *ACM Computing Surveys*, 55(12), 1–38. <https://doi.org/10.1145/3571730>
- Jia, W., & Pun, J.** (2025). Exploring EMI teachers' agency in addressing language-related challenges with ChatGPT: A multiple case study. *Computer Assisted Language Learning*, 1–29. <https://doi.org/10.1080/09588221.2025.2511061>
- Kaplan-Rakowski, R., Grotewold, K., Hartwick, P., & Papin, K.** (2023). Generative AI and Teachers' Perspectives on Its Implementation in Education. *Journal of Interactive Learning Research*, 34(2), 313–338. <https://doi.org/10.70725/815246mfssgp>
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., ... Kasneci, G.** (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274. <https://doi.org/10.1016/j.lindif.2023.102274>
- Ladson-Billings, G.** (1995). Toward a Theory of Culturally Relevant Pedagogy. *American Educational Research Journal*, 32(3), 465–491. <https://doi.org/10.3102/00028312032003465>
- Lee, S., Jeon, J., & Choe, H.** (2025). Generative AI (GenAI) and pre-service teacher agency in ELT. *ELT Journal*, 79(2), 287–296. <https://doi.org/10.1093/elt/ccaf005>
- Liu, Z.** (2025). Human-AI Co-Creation: A Framework for Collaborative Design in Intelligent Systems. *AHFE International*. <https://doi.org/https://doi.org/10.48550/arXiv.2507.17774>
- Muliastuti, L., Leiliyanti, E., Eriyani, R. N., Friyandini, R., Ariestira, A. R., & Darmawati, B.** (2025). Development of Mandar Language Teaching Materials for Language and Cultural Preservation among High School Students in West Sulawesi. *International Seminar on Humanity, Education and Language*, 1(1), 1581–1585. <https://doi.org/https://doi.org/10.21009/ishel.v1i1.62902>
- Musaeus, L. H., Petersen, M. G., & Klokmoose, C. N.** (2024). Bringing Teachers and Researchers Together through Participatory Design and Cooperative Prototyping in Computing Education. *Proceedings of the 55th ACM Technical Symposium on Computer Science Education V. 1*, 902–908. <https://doi.org/10.1145/3626252.3630796>
- Nguyen, L. V.** (2025). Culturally Responsive AI Scaffolding in Collaborative Language Education: Exploring Sociocultural Adaptations across Diverse Learning Contexts. *Journal of Posthumanism*, 5(10), 353–367. <https://doi.org/10.63332/joph.v5i10.3551>
- Nianty, D. A., Asbara, N. W., Z, N., Ridwan, R., Bahasoan, S., Rahmah, R., Khalik, A., & Salnia, S.** (2024). Perahu Sandeq: Warisan Suku Mandar yang Terancam Punah oleh Modernisasi. *Jurnal Manajemen STIE Muhammadiyah Palopo*, 10(2), 358. <https://doi.org/10.35906/jurman.v10i2.2206>
- Nicholson, R., Bartindale, T., Kharrufa, A., Kirk, D., & Walker-Gleaves, C.** (2022). Participatory Design Goes to School: Co-Teaching as a Form of Co-Design for Educational Technology. *CHI Conference on Human Factors in Computing Systems*, 1–17. <https://doi.org/10.1145/3491102.3517667>
- Raeburn, A.** (2025). Apa itu AI yang Berpusat pada Manusia? Pendekatan Empatik terhadap Teknologi. *Asana*. <https://asana.com/id/resources/what-is-human-centric-ai>
- Trang, T. T. T., & Hieu, H. N.** (2026). Representasi Kearifan Lokal Masyarakat Mandar pada Cerpen Rumah Batu Kakek Songkok Karya Lina PW. *Anafora: Jurnal Penelitian Mahasiswa Pendidikan Bahasa dan Sastra Indonesia*, 6(2), 238–248. <https://doi.org/https://doi.org/10.25134/ajpm.v6i2.325>
- Yusniawati, N. L. P. I., & Lestari, Y. B.** (2021). *The Potential of Local Culture in English Language Teaching (ELT): A Response Paper to Domination of English Material in ELT*: 2nd Annual Conference on Education and Social Science (ACCESS 2020). <https://doi.org/10.2991/assehr.k.210525.098>
- Zhai, X.** (2024). *Transforming Teachers' Roles and Agencies in the Era of Generative AI*:

*Perceptions, Acceptance, Knowledge, and
Practices* (Version 1). ArXiv
<https://doi.org/10.48550/ARXIV.2410.03018>