

Exposing the Gap: AI-Assisted AQRF Evaluation of Grade 12 English Textbooks in Indonesia's Kurikulum Merdeka for Edu-Revolution

Muhammad Azkiya Bahtsulkhoir¹, Salsabilla¹, Ananditamala Syalwa Heryana¹

¹ UIN Siber Syekh Nurjati Cirebon, Indonesia

*Corresponding author's email: muhammadazkiyabk@mail.syekhnurjati.ac.id

 <https://doi.org/10.54855/paic.2693>

®Copyright (c) 2026 Muhammad Azkiya Bahtsulkhoir, Salsabilla, Ananditamala Syalwa Heryana

Received: 30/09/2025

Revision: 15/05/2026

Accepted: 14/06/2026

Online: 23/07/2026

ABSTRACT

Keywords: Artificial Intelligence, Kurikulum Merdeka, Reading for Emotions, Aesthetic Principles

As Indonesia advances toward the Indonesia Emas 2045 vision, English education is expected to prepare learners for global collaboration. However, debates remain on whether Kurikulum Merdeka textbooks foster higher-order thinking, emotional engagement, and aesthetic literacy. While many studies analyze textbooks, few incorporate emotional and aesthetic dimensions, and even fewer explore AI as an assessor in textbook analysis. This study examines the extent to which a Grade 12 Kurikulum Merdeka English textbook develops competencies aligned with AQRF Level 4, such as independent problem solving, critical engagement, and applying knowledge in unfamiliar contexts. Situated in the context of Grade 12 English learning in Indonesia, this pilot study analyzes a narrative text, *The Breakers of Limitations* from Gurung Mali, using AI-assisted textual analysis with ChatGPT, Qwen, and Gemini. Guided by Reading for Emotions and neuroaesthetic principles with AQRF-informed prompts, the findings show cultural richness and engaging narratives but limited emotional depth, aesthetic appeal, and higher-order reasoning, highlighting the need for an AI-driven, AQRF-informed redesign.

Introduction

Indonesia Emas 2045, as the ultimate national vision, conveys much more than just a demographic concept. It is a scenario in which the majority of the population is of productive age and, as such, represents a strategic opportunity to drive national progress (Ansori, 2021; Sudarma, 2022). Nevertheless, repeated budget cuts in important areas, such as education, carried out under Inpres No. 1 *Tahun* 2025 (Republic of Indonesia, 2025), significantly jeopardize this opportunity, as they not only impede its realization but also deepen existing knowledge and skill gaps.

In this context, improving the quality of Indonesian human resources becomes a central strategic priority, particularly in the mastery of globally influential languages such as English.

CITATION | Bahtsulkhoir, M. A., Salsabilla, & Heryana, A. S. (2026). Exposing the Gap: AI-Assisted AQRF Evaluation of Grade 12 English Textbooks in Indonesia's Kurikulum Merdeka for Edu-Revolution. *Proceedings of the AsiaCALL International Conference*, 9, 40-62. ISSN: 2833-6836, ISBN: 979-8-90148-783-9. DOI: <https://doi.org/10.54855/paic.2693>

Strengthening English education not only enhances individual competitiveness in an increasingly globalized labor market but also contributes to the realization of Sustainable Development Goal 4 (SDG 4) on quality education. In today's globalized world, proficiency in a foreign language is a highly valued skill that enhances individuals' professional competitiveness and career advancement opportunities (Nguyen et al., 2020). English, in particular, plays a crucial role in science and technology, as it facilitates international communication and fosters global competitiveness, innovation, and cross-border collaboration for sustainable development (Marliana, 2024).

Moreover, as English increasingly becomes the working language of ASIA and the ASEAN region, the mobility of people and information has intensified intercultural communication. The massive use of English among native speakers creates new challenges in language education: it not only focuses on linguistic competence, but also intercultural awareness and meaningful engagement (Lian & Sussex, 2018). Despite its importance, Indonesia continues to face significant challenges in providing effective English-language education.

The 2024 EF English Proficiency Index indicates that Indonesia is ranked 80 out of 116 countries and regions, meaning that the position is significantly lower than that of other Southeast Asian countries, for instance, Malaysia and the Philippines (EF Education First, 2024), despite the various efforts and reforms in policies implemented from 2011 to 2024 to enhance English education and skills training. The poor ranking calls for immediate action on English proficiency, to be addressed through systematic and sustainable efforts. While the English language has been instrumental in the exchange of knowledge and technology, the advancement of Indonesian human resource quality is still heavily dependent on a wide range of factors, including academic, professional, and personal factors, at both the national and global levels (Gasani, 2025).

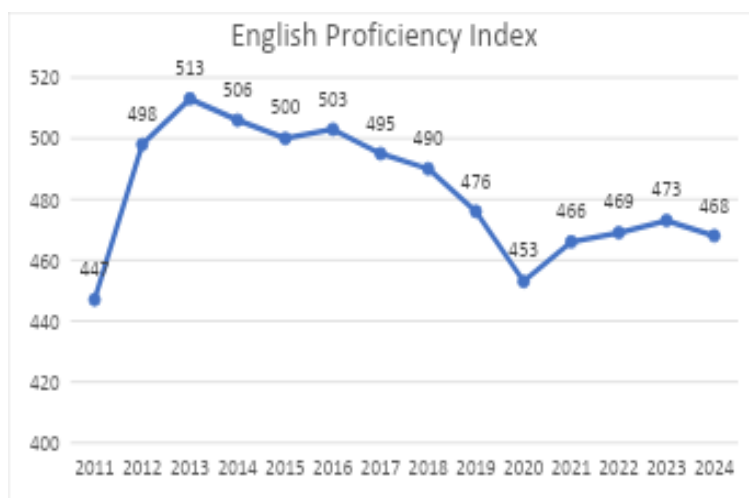


Chart 1: Graph Indonesian EPI Score 2011-2024

Besides reflecting limited proficiency, the low EPI score also leads to suspicion that primary learning resources in schools, Kurikulum Merdeka textbooks, which are considered the basic materials (Damayanti & Mukarto, 2024) and the main language input components for English as a Foreign Language (EFL) in Indonesia (Senowarsito et al., 2023), might not be of good quality. Kurikulum Merdeka, an Indonesian government initiative, is designed to promote meaningful learning that is in line with students' needs. In such a scenario, textbooks are expected not only to teach grammar and vocabulary but also to develop students' critical

thinking skills, provide meaningful text engagement, introduce them to other cultures, and help them acquire problem-solving skills for lifelong learning and global competitiveness (Wachyuniyah et al., 2024). If students lack these qualities, they will find it difficult to perform well in international assessments, participate in global discourse, or adjust to complex professional and academic environments. Therefore, the insufficiencies of these textbooks not only diminish individual learning outcomes but also, on a larger scale, pose a challenge to national objectives, potentially impeding the realization of Indonesia Emas 2045, which predicts a generation of human resources that are internationally competent, innovative, and competitive.

In this regard, the Grade 12 English readers should be compliant with Level 4 of the ASEAN Qualifications Reference Framework (AQRF), which means that they should have broad technical and theoretical knowledge, be able to adapt predictable work processes to new situations, and engage in self-direction and coordination when dealing with unfamiliar problems (ASEAN Secretariat, 2012; Komite Nasional Kualifikasi Indonesia, 2020). Achieving this level ensures that students are not only capable of sitting comfortably for national examinations but also possess the necessary skills to compete at and succeed in wider regional and international contexts, as the AQRF is one of the main components in facilitating comparability between standards in different ASEAN member states (Kuroda et al., 2018; Carvajal & Vicente, 2024). As the curriculum revamp is Indonesia's national priority, Indonesia is not only creating a more robust local education system but is also playing its part in the realization of ASEAN's five-year plan for a digitally savvy and vibrant region (Sudimantara et al., 2025).

The analysis mainly refers to Piaget's (1969) cognitive developmental theory and particularly highlights the formal operational stage when students become capable of abstract and critical thinking beyond concrete reasoning. A sort of cognitive reproductive failure may occur when the exposure they get is not in line with their developmental needs, as cognition integrates perception, emotion, and logic into what Piaget calls operations, mental processes that not only systematize understanding but also regulate perception in life. This cognitive change is very much in line with Zull's (2011) argument that learning is initially based on the brain's concrete structures and functions and then moves to the abstract realm of the brain. By sensing, recording, and reproducing experiences, the brain increases its potential to think, decide, and act, thereby becoming a reflective mind capable of awareness and meaning-making. Such a process is continuous and represents a learning cycle (Dewey, 1938) which eventually becomes the basis for elevated levels of cognitive abstraction. At the level of formal operations, learners should be able to attain *conscientização*, the stage of critical consciousness that empowers people to see themselves as subjects with the capacity to change the world. Hence, English textbooks ought to feature generative themes that are essentially based on real and relatable human experiences, thus being in tune with Wirawan's (2025) BIMA (Bridging Intelligence, Mindfulness, and Awareness) model, which refers to Regenerative Learning as a process through which learning continuously renews awareness and compassion via conscious reflection and transformative experience. The presence of such elements in the curriculum fosters awareness, freedom, and a sense of responsibility for social and human transformation (Freire, 1970). When put together, these theoretical perspectives conform with the AQRF Level 4 standards that highlight the importance of knowledge adaptation, independent learning, and self-directed problem-solving. This holistic view offered by the integrated lens helps to ascertain if English textbooks not merely facilitate the acquisition of higher-order thinking and emotional literacy but also foster the development of mindful awareness, which, among others, are the main prerequisites for the realization of the globally competent learners envisioned by

Indonesia Emas 2045. The metaphor of 1945 Indonesia's independence not only as the liberation of the body and nation, but also of the mind through education, which subsequently becomes the vehicle of freedom, evolving from simple sovereignty to intellectual and moral maturity, eloquently captures the essence of the 2045 centennial vision.

Although the Kurikulum Merdeka was hailed as a transformative reform toward competency-based and student-centered learning, initial analysis of the Grade 12 English textbook, the core material at the final stage of secondary education, exposes some deficiencies in the competencies expected to be achieved at Level 4 of the ASEAN Qualifications Reference Framework. At this level, learners should already demonstrate broad technical and theoretical knowledge, adaptability in predictable yet changing processes, and self-direction in proactively addressing unfamiliar problems. However, the textbook's content seems deficient in consistently fostering these abilities, which also echoes national concerns about Indonesia's ranking in the 2024 EF EPI and the risk of assuming that textbook availability ensures competency. This calls for an urgent need to develop a far stronger and more neutral evaluation mechanism for the English textbooks. In this respect, AI, as a promising evaluator, is expected to revolutionize the educational system through adaptive, real-time evaluations of higher-order skills enabled by technologies such as machine learning and natural language processing. It also offers dynamic, data-driven, reproducible assessments that complement human judgment, aligning with policy aspirations and actual learning outcomes.

AI development is a paradigm shift—a movement with the power to fundamentally alter many aspects of life, from the way we live together in society to how our educational constructs assess and evaluate. It is obvious that the rise of AI is a game-changing phenomenon that has many advantages in various fields (Pham, 2025). Generally speaking, AI refers broadly to computational systems capable of performing tasks that typically require human intelligence, and AI applications within education have grown increasingly prevalent in adaptive learning, automated feedback, and large-scale assessment as paths toward streamlining evaluation processes. As an evaluator itself, AI can process volumes of textually discovered patterns therein and apply consistent criteria—all of which minimize human subjectivity and potential bias. Later developments within the domain of NLP allow models such as ChatGPT, Qwen, and Gemini to progress beyond merely grammatical surface-level text analysis to examine deeper levels of texts regarding cognitive demands, emotional resonance, and aesthetic quality. This puts them in a position to check whether actual practice in English textbooks inspires the competencies, rather than being superficially assured by simple checks for content coverage or alignment with learning outcomes.

In this study, AI serves as an advanced tool in the multi-dimensional assessment of the Grade 12 Kurikulum Merdeka English textbook. Neither the Reading for Emotions model by Lian (2017) nor the aesthetic principles by Ramachandran & Hirstein (1999) are merely stylistic or analytical models; they are cognitive-affective literacy frameworks that center reading as an emotional and intellectual journey. According to Ania Lian (2017), the Reading for Emotions model considers texts as emotive constructs created with a view to engendering affect in the reader. Here, comprehension is less dependent on the text than upon the dynamic "conversations" and purposes that the reader brings into engagement, which inherently render reading a critical and creative activity. In this light, transformative learning can be seen as reflection upon those emotional and intellectual dialogues. The AI models, ChatGPT, Qwen, and Gemini, for instance, identify emotional cues, aesthetic patterns, and cognitive complexity that distinguish passive and active texts, or texts which reproduce or stir emotional and intellectual awareness. Reading then becomes an act of exploration of the "learning organ" and shows exactly how the journey from the brain to the mind is facilitated through emotionally

resonant and cognitively captivating material.

Coupled with AI-assisted assessment, this paper is expected to gauge the extent to which the Grade 12 Kurikulum Merdeka English textbook develops the competencies required for AQRF Level 4, such as solving problems independently, critical engagement, and applying knowledge in unfamiliar contexts. While much research is conducted on textbooks, few studies incorporate emotional and aesthetic dimensions, and even fewer explore the use of AI as a potential assessor in textbook analysis. It is envisaged that the findings will offer actionable insights for policymakers, curriculum developers, and educators in redesigning learning materials that develop emotional engagement, aesthetic literacy, and higher-order thinking. Emphasizing transformative learning as its core strategy, Indonesia can thus strengthen its education system and at the same time contribute to ASEAN's vision of a digitally advanced and dynamic region (Sudimantara et al., 2025). The approaches highlighted here provide ways for English education to become a driver of literacy, innovation, and collaboration in support of the realization of Indonesia Emas 2045.

Literature review

Artificial Intelligence has now emerged as one of the most transformative forces across a wide array of domains, bringing new dimensions in information processing, insight development, and human decision-making. It is defined as any computational system that can do things which, if done by humans, would be said to require intelligence: reasoning, learning, and problem solving (Russell & Norvig, 2021). AI has thus far illustrated substantial promise within education. In particular, it has the potential to facilitate personalized learning, automate administrative processes, and create adaptive feedback systems responsive to the needs of individual learners (Ji et al., 2023). Unlike other conventional digital tools, AI systems analyze complex patterns in data, automate routine tasks, and provide scalable instructional support—all arguments that make them promising solutions for addressing challenges associated with large class sizes, diverse learning styles, and increasing efficiency demands on teaching and assessment. However, scholars also advise that deploying AI within education raises concerns about fairness, bias, accountability, and the irreplaceable value of human judgment within contexts requiring empathy, ethical reasoning, and creativity (Selwyn, 2019; Williamson & Eynon, 2020). These debates highlight the need to conceptualize AI not as a substitute for educators but as a complementary partner in enhancing pedagogical practice.

Compared to other fields, the educational field illustrates a rather specialized area of implementation in which AI can more easily be used to speed up the process. Recent studies have underlined the increasing role of AI as an assessor in educational contexts: both the potential benefits and limitations are documented. Trikoili et al. (2025) reported significant differences between ChatGPT and human raters. AI provided quick first-order assessments, whereas educators contributed interpretative insights; combining AI tools with human judgment may optimize critical thinking assessment. Correspondingly, Karademir and Alper (2024) noticed that AI-based assessment systems offer consistency, automated grading, personalized feedback, and data-driven insights that enable teachers to devote much more time to teaching strategies and student support. However, individual differences and nuanced interpretations provide evidence that AI cannot replace human assessors; rather, this tool should enhance the assessment process and support educators in refining learning results in greater detail (Celik et al., 2022; Mahligawati et al., 2023).

Apart from assessment, AI is increasingly explored as a collaborative and creative partner in

education. Atchley et al. (2024) highlight AI's neutral evaluation in group work and the preparation of students for professional settings that expect cooperation with AI. Urmeneta & Romero (2025) underline the potential of AI to support students' creativity by giving feedback, evaluating creative products, and predicting achievement. Further, work on human-AI agency suggests that integrating AI in education raises questions about a balanced orientation between technological capabilities and human critical engagement. Mouta et al. (2025) and Krakowski (2025) make quite strong arguments that human agency and imagination are central in shaping how AI technologies are applied with a view to preserving collective autonomy, ethical considerations, and aspirational goals. Taken together, what these studies provide is a theoretical basis for the use of AI in textbook assessment, underlining its potential for the analysis of cognitive, emotional, and creative dimensions while retaining the place of human oversight.

Yet, despite these advances, several gaps are still evident in existing research. Whereas much research has been conducted on the role of AI in assessment, creativity support, and collaborative learning, far fewer studies have considered its role as a systematic assessor of national curriculum textbooks against international standards such as the ASEAN Qualifications Reference Framework. Most past work has focused either on single-model AI or on general applications of AI, without considering the comparative interpretive capabilities of multiple chatbots. To the best of our knowledge, very few, if any, studies have compared how different AI systems, such as ChatGPT, Qwen, and Gemini, analyze the same educational material and assess its alignment with competency frameworks. Further, when compared with classroom assessment or adaptive learning, textbook evaluation remains a relatively underexplored research focus, despite being foundational in shaping students' cognitive, emotional, and aesthetic engagement. This study thus responds to two large gaps: one in comparing multiple AI evaluators and another in investigating AI as a critical partner in reviewing national curriculum textbooks, beyond efficiency toward quality, inclusivity, and international compatibility.

The present research adds to these pedagogical and technological considerations by incorporating the Reading for Emotions framework (Lian, 2017) and aesthetic principles (Ramachandran & Hirstein, 1999) as cognitive-literary models focused on the affective dimensions of text. These frameworks place the act of reading as an emotional and perceptual journey whereby meaning is constructed in the interplay among feeling, thinking, and perception—a dynamic process that bridges emotional resonance with cognitive understanding. Instead of being treated as a passive linguistic entity, texts are viewed as living constructs that act to evoke and organize emotions, thus nurturing habitual cognitive engagement through images and words. This perspective echoes Zull's (2011) contention that learning runs parallel with the basic operations of all functioning nervous systems, where every linguistic image fires neural pathways that convert sensory experience into emotion, reasoning, and awareness. Integrating these frameworks, AI-assisted evaluation goes beyond linguistic accuracy to understand how language activates emotional, aesthetic, and cognitive circuits in learners—to engender holistic intelligence and reflective awareness aligned with Kurikulum Merdeka and the transformative vision of Indonesia Emas 2045.

Research Questions

Taking as its base previous studies on AI-assisted evaluation (Trikoili et al., 2025; Karademir & Alper, 2024; Urmeneta & Romero, 2025) and debates regarding human–AI agency in education (Mouta et al., 2025; Krakowski, 2025), the present study explores questions about the potential of AI in textbook assessment by asking:

1. To what extent can the AI models of ChatGPT, Qwen, and Gemini assess the Grade 12 Kurikulum Merdeka English textbook, most specifically the text "The Breakers of Limitations from Gurung Mali" (Hardini et al., 2022, pp. 13–17), in reference to AQRF Level 4 competencies in higher-order thinking, emotional engagement, and aesthetic literacy?
2. What does the AI-assisted evaluation reveal about textbook strengths and weaknesses, and in what ways can AI complement or augment traditional human assessment to identify gaps and inform improvement in learning materials in English?

Methods

Pedagogical Setting & Participants

This research is situated within the context of Grade 12 English learning in Indonesia's *Kurikulum Merdeka*. The study focuses on a single narrative text, *The Breakers of Limitations* by Gurung Mali (Hardini et al., 2022, pp. 13–17), drawn from the official Grade 12 English textbook. Designed as a pilot study, the text serves as the primary unit of analysis representing instructional material intended to develop students' cognitive, emotional, and aesthetic competencies.

The analytical process employed an AI-assisted textual analysis using multiple language models (ChatGPT, Gemini, and Qwen). These models were utilized to examine the narrative for emotional and aesthetic features in relation to the selected theoretical frameworks. The use of multiple AI tools allowed for comparative pattern identification across models, highlighting convergences and divergences in analytical outputs. This approach aligns with the learning expectations of the final stage of secondary education, where learners are assessed based on higher-order competencies outlined in AQRF Level 4, including problem-solving, critical thinking, and the application of knowledge in unfamiliar contexts.

Given the pilot nature of the study and its focus on a single narrative text, the findings should be interpreted with caution. While the selected text is treated as a representative instructional unit, variations in genre, thematic emphasis, and pedagogical objectives across other sections of the textbook may yield different cognitive, emotional, or aesthetic patterns. Accordingly, the results are exploratory and intended to inform future, broader-scale analyses rather than support generalization across the entire textbook.

Design of the Study

The study adopted a qualitative design, integrating thematic content analysis of selected text material. Qualitative research, according to Tenny et al. (2025), offers detailed insight into complex phenomena in real-life situations. Content analysis, on the other hand, entails an objective and systematized method of studying written data, as observed by Bengtsson (2016). The Reading for Emotions framework by Lian (2017), aesthetic principles by Ramachandran & Hirstein (1999), and AQRF Level 4 standards were combined in this analysis, allowing the current study to explore emotional cues, cognitive complexity, and aesthetic features within the text.

Data collection & analysis

AI models employed in the analysis include ChatGPT, Qwen, and Gemini. The analytical process began with AI-assisted textual analysis, in which prompt engineering was guided by the Reading for Emotions framework and aesthetic principles, with AQRF Level 4

competencies incorporated through few-shot learning supported by theoretical explanations and illustrative examples. Throughout this stage, the AI systematically examined multiple dimensions of the text, including emotional and cognitive content, aesthetic richness relative to standard expectations, and textual alignment with AQRF Level 4 competencies.

To enhance analytical reliability, multiple AI models were applied to the same textual data, allowing for comparative examination of convergent and divergent patterns across outputs. This cross-model comparison functioned as an internal validation mechanism, strengthening the consistency of AI-assisted findings. The resulting analysis offers an exploratory evaluation of the extent to which the selected text supports higher-order thinking, emotional literacy, and aesthetic awareness in alignment with national and regional educational frameworks, while acknowledging the need for further validation through expanded datasets or complementary evaluative approaches in future research.

Analytical Prompt Structure

The AI-assisted analysis in this study was conducted using a structured analytical prompt designed to operationalize the Reading for Emotions and Reading for Aesthetics frameworks in a systematic and reproducible manner. The prompt instructed the model to organize its analysis into clearly defined conceptual categories and to present findings in tabular form to support analytical clarity. Specifically, the model was guided to identify and interpret six core components of Reading for Emotions Focus, Disturbance, Dialogue, Development, and Moral before extending the analysis to aesthetic perception using key principles such as Peak Shift, Grouping, Contrast, Symmetry, Perceptual Problem Solving, Metaphor, and Generic Viewpoint. The same prompt was applied consistently to ensure structured interpretation of the narrative text and to standardize analytical output across AI-generated responses.

The full prompt used in the AI-assisted textual analysis was as follows: “Create a table presenting an analysis of the text *The Breakers* by Gurung Mali that has been provided. Organize the analysis according to the six main components of the Reading for Emotions framework (Focus, Disturbance, Dialogue, Development, and Moral). After presenting the emotional analysis, examine the structure of both Reading for Emotions and Reading for Aesthetics. For the aesthetic analysis, evaluate the text using the following principles: Peak Shift, Grouping, Contrast, Symmetry, Perceptual Problem Solving, Metaphor, and Generic Viewpoint. Provide clear analytical explanations for each category and present the results in a structured and organized format.”

Results/Findings and Discussion

The Reading for Emotions framework highlights the role of emotion as a critical lens in analyzing narrative texts and helping learners not only comprehend the story but also experience and recognize emotional changes throughout the reading process. Based on the traditional structural elements of narrative, which include Orientation, Complication, Resolution, and Coda (Cadwell & White, 2017), Lian extended the framework by adding two more, Dialogue and Development, to present six stages: Focus, Disturbance, Dialogue, Development, Resolution, and Moral. This encompasses the opening emotional tone, shifts in affect, responses through internal reflection or character interaction, unfolding events that sustain engagement, narrative resolution, and the underlying moral or reflection embedded in the text. As such, this approach allows for a nuanced understanding of how narratives guide readers’ emotional journey and engagement with content.

This study uses AI as a high-end assessor. Applying Lian's integrated Reading for Emotions and neuroaesthetic framework Ramachandran & Hirstein, 1999, it assesses the Grade 12 Kurikulum Merdeka English textbook, specifically the story *The Breakers of Limitations* from Gurung Mali, pages 13 to 17. The AI models, ChatGPT, Qwen, and Gemini, systematically scan for emotional cues, aesthetic features, and cognitive demands in the text. These are then mapped onto AQRF Level 4 competencies that expect higher-order thinking, independent problem-solving, and critical engagement in unfamiliar contexts. By capitalizing on AI's ability to process massive volumes of text with consistent criteria, the study explores whether the narrative develops emotional literacy, aesthetic awareness, and cognitive sophistication and points toward the gaps and strengths that would have been overlooked in the traditional human examination. This integration of AI ensures the assessment is accurate and multidimensional, reinforcing the transformational and critical learning goals of Kurikulum Merdeka.

Table 1.

Full Text of *The Breakers of Limitations* from *Gurung Mali* (Grade 12 Kurikulum Merdeka, pp. 13–17)

The Breakers of Limitations from Gurung Mali



One night, in the village of Gurung Mali in Sintang district of West Kalimantan, the night was colder and darker. There was no light at all. The enchanting beauty of Gurung Mali village during the day felt short and the night was very long. People couldn't do anything other than go to bed, including a man named Nayau. Nayau often contemplated what he could do so that the beauty of his village could be seen even longer. This would make people be able to do more activities at night without any limitation.

The existence of electricity in their village was impossible to wait for because the location of the village was very far from the government's electricity installation facilities. Nevertheless, it did not make Nayau be in despair. When other people could only accept the situation, Nayau began to gather his friends to do something for the sake of the light for their village and for breaking up the limitation they had.

A faint conversation was heard among a group of men and women. They gathered before the campfire while roasting corn and drinking hot tea.

"Your idea is impossible Nayau", said Paoi, "How can we build a micro hydropower plant? How much money must it cost to build it?"

"Yes Nayau, do you have that much money?", Nico asked.

"That's right, moreover the distance of our village from the Filing river is also very far, Nayau" Paoi scolded.

Nayau smiled and said, "Yes, I don't have much money, friends. But if we work together,

nothing is impossible. This is for our goodness. Our village can be bright and beautiful at night and we can also do more productive activities at night without any limitations like now”

“If we don’t have money, how can we build a ditch to drain water from the Filing river Nayau?” asked Kosa, a woman engrossed in adjusting the position of

the wood in the pit.

“We can try to use wood as it is cheaper”, said Nayau.

“How about using diesel fuel to drive the wheel?” asked Sinko.

“If we use diesel, we have to provide funds every day to buy the diesel and it can also pollute the air into our environment. If we use hydropower, it’s more efficient and more environmentally friendly”, Nayau explained.

“But using wood is still expensive, Nayau, because the distance is very far. We need to buy a lot of wood to build the trench”, said Apoi.

“Building the windmill also requires a lot of money, Nayau”, said Noh.

“We can borrow the material first, then we can pay it in installments”, said Nayau.

“Nowadays, there is no material owner who wants to give their goods without any collateral”, said Dasen.

“I am ready to be a guarantor for the sake of electricity in our village”, said Nayau.

“Wow, that’s very unreasonable. I don’t agree. I’m afraid the project will fail. There’s no electricity and money will be gone. For the worst scenario, we still need to pay for the material”, said Miso, who had been silent for a long time.

Then some people left the group. Only a few remained, they were still sitting quietly.

“If we don’t try to change our condition, who else will change it for us?” Nayau said quietly.

“We can’t just stand still, waiting for the electricity to come from the government and we don’t know when it will come”, Nayau added.

“That’s right”, said Apoi who was still sitting closest to the firebox.

“But I can’t make it happen alone. Together is stronger than alone, right?” said Nayau.

“Yeah, let’s do it together, I’m ready to help”, said Supardi.

“We can also help, the women’s troops are ready”, said Nino, the youngest woman among the others, “Come on, we can do it together”, Hisako said, clenching his fists.

On the next day, they came to the materials shop to discuss the matter. After a few bargaining and such, they made an agreement. They had a deal. Nayau and several supportive friends such as Apoi, Sinko, Hisako, Noh, Dasen, Supardi, Mio, Nino and several other women began working together to build a dam and a mill house along with the water channel in the form of a ditch that reached 200 meters in length and 40 cm in depth. They worked collectively, piled sandbags for the dam, covered the mill house with tarpaulin and covered the generator with zinc. Everything was done manually. Every day, their profusely sweat was not felt as a barrier to their passions to change the condition of their village. They were eager to make their village enlightened and eliminate the limitations of their activities at night.

After three months, the micro hydropower plant was finally completed. Finally, the houses where they lived were in joy at night. Those who previously doubted the value of hard work,

cooperation and care for the environment then joined in cheers. Villagers were then more enthusiastic about protecting the forest and the surrounding environment to keep the water supplies remaining available. Nayau smiled broadly and was proud of his friends.” The beauty of Gurung Mali village will stay forever and people will be more prosperous since they are more productive now”, Nayau whispered “Togetherness can break through limitations”

Table 2.

ChatGPT Analysis of *The Breakers of Limitations from Gurung Mali* Based on AQRF Level 4: Emotional, Aesthetic, and Cognitive Dimensions

Story Segments	Analysis
“One night, in the village of Gurung Mali... There was no light at all.” “Nayau often contemplated what he could do so that the beauty of his village could be seen even longer.”	Both segments share the emotional stage of Focus but emphasize different aspects. The first uses Peak shift, Contrast, and Symmetry to portray hopeless darkness, remaining descriptive (AQRF 2-3). The second shifts to reflective hope through Perceptual problem-solving and a Generic viewpoint, but stays abstract, fitting only pre-AQRF 4.
“Your idea is impossible, Nayau...” “I’m afraid the project will fail.”	The emotional stage reaches Disturbance, with skepticism and fear creating strong tension. Through Contrast, Grouping, and Metaphor, the conflict is intensified, making this segment highly engaging and aligned with AQRF 4 due to its collaborative problem-solving and adaptability.
“If we work together, nothing is impossible... Together is stronger than alone.”	It reflects Dialog on optimism as an emotional stage, supported by Symmetry and a Generic viewpoint (aesthetic principles), but its cliché phrasing limits emotional originality, aligning more with AQRF 3 (basic persuasion/communication).
“Yeah, let’s do it together, I’m ready to help”, said Supardi... “We can also help, the women’s troops are ready.” “They began working together to build a dam... Every day, their profuse sweat was not felt as a barrier.”	All sentences reflect the Development stage, showing a shift from doubt to collective support. The first and second dialogue uses Perceptual problem-solving and Peak shift to unify voices into action, aligning with AQRF 4. In contrast, the last dialogue relies on Symmetry and Contrast but remains repetitive and emotionally flat, fitting closer to AQRF 3.
“After three months, the micro hydropower plant was finally completed. Finally, the houses... were in joy at night.”	This sentence represents a Resolution segment which employs Contrast and an exaggerated Peak shift of joy, but feels too smooth and skips the struggle, making it only partially fit for AQRF 4.
“The beauty of Gurung Mali village will stay forever... Togetherness can break through limitations	This sentence reflects a Moral that works through Metaphor and a Generic viewpoint, appealing logically rather than emotionally, and aligns again with AQRF 3 instead of demonstrating higher-level competencies.

ChatGPT highlights that the story is strong at establishing atmosphere, especially effective during the campfire dialogue, where skepticism, fear, and negotiation create vivid emotional tension. Aesthetic principles such as contrast and peak shift are well used, reinforcing the

symbolism of darkness and light. However, other sections lack depth; the introduction and resolution are more descriptive than evocative, and much of the dialogue feels cliché, reducing originality. The construction phase reads like a report with no struggles or setbacks, weakening emotional impact. Thus, while some parts reflect Level 4 qualities (collaborative reasoning and decision-making), overall the text remains closer to Level 3, as it does not consistently show complex problem-solving or advanced reflection.

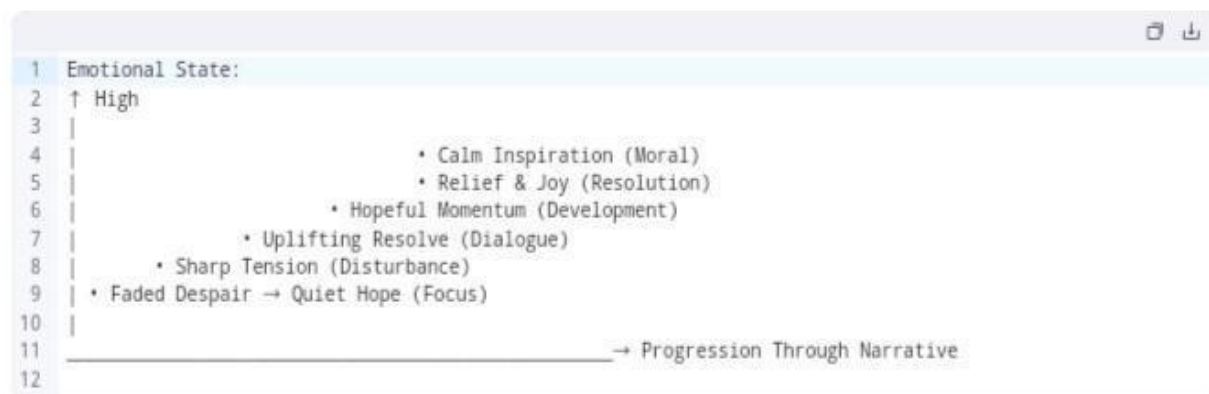


Chart 2. Emotional State on the Analysis by ChatGPT

This graph reflects the story's emotional rhythm: it opens with a superficially dark but underdeveloped atmosphere that quickly shifts to Nayau's quiet hope, lacking deep communal emotion. The peak emotional intensity occurs during the Disturbance, where genuine skepticism and financial fear create authentic tension. While Dialogue and Development raise spirits through collective resolve and unity, these stages lean on familiar tropes and lack visceral struggle or personal setbacks, softening their emotional impact. The Resolution delivers satisfying relief but feels overly smooth, missing a hard-won climax. Finally, the Moral lands as thoughtful and principled, logically resonant, yet emotionally restrained rather than deeply stirring.

Table 2.

Qwen Analysis of *The Breakers of Limitations from Gurung Mali* Based on AQRF Level 4: Emotional, Aesthetic, and Cognitive Dimensions

Story Segments	Analysis
"One night, in the village of Gurung Mali... There was no light at all."	Emotions: Focus. Unrealistic exaggeration ("no light") ignores real-world solutions, portraying helplessness. Peak shift is overdone without artistic subtlety; contrast (light/dark) is a tired trope; Symmetry is simplistic syntax, not intentional craft. Scene-setting fails, emotionally vacant, tells not shows. Lacks Perceptual Problem Solving. Rated AQRF 1-2: passive observation without technical insight.
"Nayau often contemplated what he could do so that the beauty of his village could be seen even longer."	Emotions: Shifts from Focus to Disturbance. Poetic wish mistaken for planning, ignores technical requirements of micro-hydro development. Absent Perceptual Problem Solving; generic viewpoint lacks specificity. Emotionally inert, abstract longing without concrete imagery. Rated pre-AQRF 4: private musing without coordination, skill application, or communal reasoning.
Campfire discussion: "Your idea is impossible, Nayau"... "I'm afraid the	Emotions: Shifts from Disturbance to Dialog. Dialogue fixates on money, ignoring real technical issues (water

project will fail.”	flow, maintenance), plot convenience over authenticity. Contrast is binary (optimism/fear), not nuanced; Grouping flattens voices into a single obstacle; campfire as Metaphor is overused and unearned. Lacks organic tension. Rated below AQRf 4: no risk analysis, feasibility study, or collaborative Perceptual Problem Solving.
Nayau: “If we work together, nothing is impossible... Together is stronger than alone.”	Emotions: Shifts from Dialog to Development. Cliche dialogue lacking emotional authenticity, reads like a motivational poster. Symmetry becomes monotony through repetitive short sentences; Generic Viewpoint reflects intellectual laziness. Manipulative rather than moving. Rated AQRf 3: basic communication without specialized knowledge, critical analysis, or adaptive strategy.
“Yeah, let’s do it together...” (Supardi, Nino, Hisako)	Emotions: Development. Superficial inclusivity, enthusiasm without defined roles for women/youth. Peak shift is unearned; emotional rise lacks buildup. Absent Perceptual Problem Solving, unity is declared, not developed through struggle. Artificial turning point lacking credibility. Rated superficial AQRf 4: collective decision-making, asserted but not demonstrated with role delegation or accountability.
“They began working together to build a dam... Every day, their profuse sweat was not felt as a barrier.”	Emotions: Development. Romanticizes manual labor while ignoring engineering realities (reinforced structures, turbines, wiring, maintenance). Repetition of “working together” functions as propaganda, not rhythm; Contrast (struggle/spirit) is told, not shown. Zero sensory detail, no blisters, mud, or frustration. Narratively inert. Rated AQRf 3: task execution without troubleshooting, innovation, or reflective Perceptual Problem Solving.
After three months... houses were in joy at night.	Emotion: Resolution. Overly smooth resolution, ignores post-launch failures, repairs, or dry season challenges. Symmetry creates narrative balance by mirroring the opening darkness with light/joy, offering closure. But emotion remains vague, “joy” is abstract, lacking sensory specificity (Isolation of feeling without texture). Rated AQRf 2-3: outcome described without evidence of applied technical knowledge, system evaluation, or critical reflection required for AQRf 4.
Togetherness can break through limitations.	Emotion: Moral. Metaphor links “togetherness” (collective action, trust) to overcoming “limitations” (darkness, poverty, isolation). Memorable but oversimplified, ignores need for external expertise, spare parts, or ongoing support. Emotional resonance lacks nuance; inspirational but not reflective. Rated AQRf 2: philosophical affirmation without critical analysis, trade-off evaluation, or evidence of autonomous Perceptual Problem Solving expected at AQRf 4.

According to Qwen’s analysis, the story is emotionally hollow, as the villagers are presented as silent props with little inner life, the conflicts feel staged, and the dialogue relies on overused motivational clichés rather than authentic human expression. Aesthetically, devices such as

peak shift, contrast, and symmetry are employed in simplistic and repetitive ways, producing flat scene-setting and artificial turning points instead of nuanced or credible development; unity is declared abruptly, while the construction phase reads like a sterile report that erases real struggle and complexity. Because of these weaknesses, the text does not meet AQRF Level 4, as it fails to demonstrate advanced problem-solving, adaptive strategies, or application of specialized knowledge, and instead reflects only the descriptive and basic communicative qualities associated with lower levels such as AQRF 2–3.

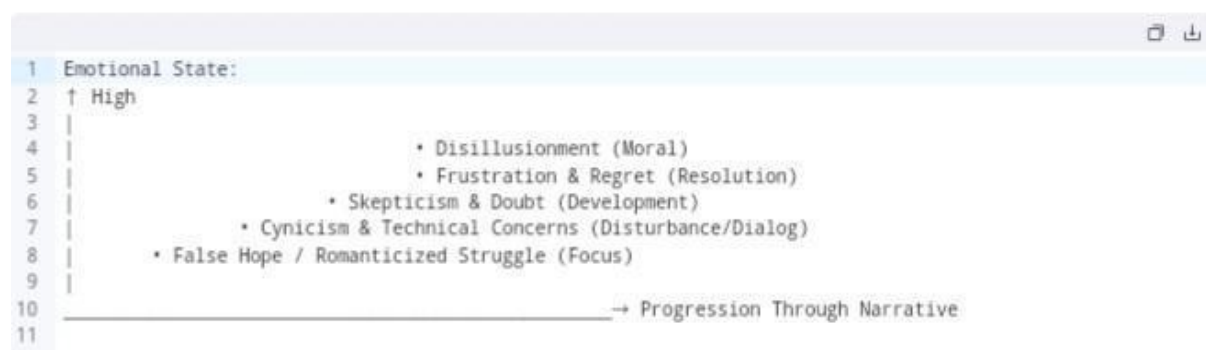


Chart 3. Emotional State on the Analysis by Qwen

This graph maps the emotional arc not as triumphant, but as a descent from romanticized idealism into sobering realism where enthusiasm collapses under the weight of unaddressed engineering, maintenance, and equity gaps. The “inspiration” promised by the moral is undercut by the absence of training, spare parts, or contingency planning, revealing a narrative more aligned with motivational clichés than sustainable development practice. The resolution’s smooth success ignores inevitable breakdowns, while the dialogue’s focus on money overlooks water flow, turbine specs, or community capacity exposing plot convenience over authenticity. Even the inclusive language around women and youth rings hollow without defined roles, skills transfer, or long-term ownership. Ultimately, the story’s emotional high point is an illusion its true legacy is not inspiration, but a cautionary tale about the cost of ignoring technical reality in favor of feel-good symbolism..

Table 4.

Gemini Analysis of *The Breakers of Limitations from Gurung Mali* Based on AQRF Level 4: Emotional, Aesthetic, and Cognitive Dimensions

Story Segments	Analysis
"One night, in the village of Gurung Mali... The existence of electricity... did not make Nayau be in despair."	Emotional stage: Focus. The stage establishes a setting of darkness where the emotions of despair are countered by Nayau's personal hope. The narrative uses Peak Shift to exaggerate the village's isolation, making the lack of light feel more intense. This is not fit for AQRF 4 because it is primarily descriptive and does not yet demonstrate the technical depth or self-directed problem-solving required by the framework.
"There was a faintly heard conversation... "We can try to use wood... "I am ready to be a guarantor... "Then some people left the group."	Emotional stage: The Core Conflict (Disturbance & Dialog). In this stage of Disturbance and Dialog, skepticism and risk create the story's strongest tension. Perceptual problem solving is engaged as characters weigh financial and logistical alternatives, driven by Nayau's moral courage. This is a strong fit

	for AQRF 4, as it explicitly demonstrates the application of broad technical knowledge and critical adaptation to solve unfamiliar community issues.
"If we don't try to change our condition... "We can also help, the women's troops are ready",... On the next day, they came to the materials shop... profusely sweat was not felt as a barrier."	Emotional stage: Collective Action (Development). The Development stage transitions to emotions of unity and commitment through sustained collective labor. The narrative uses Grouping to link individual tasks like digging a 200m ditch into a singular act of perseverance. This is only a partial fit for AQRF 4; while it shows group autonomy, it lacks the detailed technical disturbance needed to prove high-level, self-directed skill application during execution.
"After three months, the micro hydropower plant was finally completed. Finally, the houses... Those who previously doubted... then joined in cheers."	Emotional stage: The Turning Point (Resolution). The Resolution concludes the external conflict, bringing emotions of joy and vindication as light returns. The use of Contrast between the initial darkness and the final cheers provides a satisfying aesthetic closure. This is not fit for AQRF 4 because the framework focuses on the complex process of using advanced skills rather than just the successful achievement of an outcome.
"The beauty of Gurung Mali village... "Togetherness can break through limitations"	Emotional stage: Lasting Impact (Moral). The Moral stage emphasizes inspiration and environmental responsibility, using the metaphor of "togetherness" breaking through limitations. This serves to link the specific project to universal values and prosperity. This is not fit for AQRF 4, as it represents a general value orientation of Level 3 rather than the demonstration of complex technical competencies.

Gemini similarly praises the campfire dialogue as the strongest segment, noting how it presents realistic challenges- financial, environmental, and logistical- that engage critical thinking and problem-solving. The use of grouping and contrast strengthens the dramatic effect, and Nayau's responses show resilience and creativity. Yet, Gemini also stresses that the later stages of the story are weak: the three-month construction is oversimplified, perseverance feels unearned, and the resolution lacks dramatic climax. The closing moral, though inspirational, is predictable and generic. Therefore, the story illustrates moments of AQRF 4 potential in its conflict, but the lack of detailed struggles and complexity in execution reduces it largely to AQRF 3.

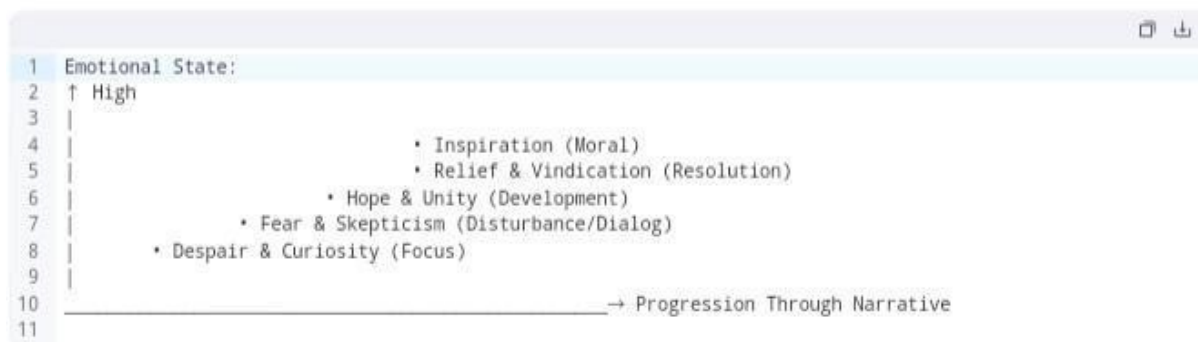


Chart 4. Emotional State on the Analysis by Gemini

This graph maps Nayau's journey from initial despair under darkness into inspired leadership, where his personal hope transforms collective fear into unified action culminating in relief, vindication, and lasting moral inspiration through environmental and communal triumph. The emotional low of skepticism during the dialogue phase represents the story's pivotal tension, overcome not by force but by Nayau's moral courage and personal guarantee. This shift catalyzes deep commitment, turning abstract hope into tangible perseverance during the long construction phase. Ultimately, the light they restore symbolizes both physical change and internal transformation validating their risk and reinforcing shared responsibility for the future.

The evaluation highlights that certain sections of the story show potential alignment with AQRF Level 4, particularly in the campfire dialogue. At this stage, the characters are confronted with skepticism, fear, and negotiation, which forces them to weigh different options, anticipate risks, and coordinate collective decisions in uncertain conditions. These moments demonstrate adaptive processes within situations that are generally predictable but subject to change, precisely what Level 4 requires. The exchange around the fire suggests that the characters are not merely acting out routine behavior but engaging in reasoning that involves technical and theoretical understanding at a general level. In this sense, the conflict segment reveals glimpses of self-direction, coordination, and problem-solving beyond the descriptive surface.

Nevertheless, the rest of the narrative falls short of sustaining these qualities. The construction phase, spanning three months, is narrated almost like a linear report, devoid of setbacks, errors, or real challenges. In an authentic Level 4 context, such a process would inevitably involve miscalculations, shifting priorities, or unforeseen obstacles that require adjustment and problem-solving. By omitting these dynamics, the story misses the opportunity to portray adaptive strategies or the application of broader theoretical knowledge in practice. Instead, it settles into descriptive sequencing that resembles AQRF Level 3, where problem-solving is straightforward and bound within routine conditions, without evidence of resilience or reflective adaptation.

From an emotional and aesthetic perspective, the text employs principles such as contrast, symmetry, and peak shift, yet in a superficial and repetitive manner that limits psychological depth and reflection. From the perspective of Paulo Freire's critical pedagogy, this resembles a "banking" approach to education, where the audience passively receives content without engaging in critical thought. Just as motivational clichés fail to provoke genuine reflection, the narrative does not foster critical consciousness (*conscientização*) or ethical awareness. In contrast, storytelling that integrates emotional tension and struggle meaningfully mirrors Freirean pedagogy by encouraging active engagement, dialogue, and reflection, empowering the audience to critically analyze experiences rather than accepting surface-level lessons.

Taken together, the story demonstrates only partial features of Level 4, concentrated mainly in the initial conflict, while the later phases remain underdeveloped and descriptive. To truly reach Level 4, the narrative would need to show more sustained adaptation, broad guidance, and coordination in overcoming unfamiliar challenges, while weaving in theoretical insights and emotional depth. Without these elements, the story remains closer to AQRF Level 3, which reflects basic communication and routine-level problem-solving rather than advanced reasoning and adaptation. In short, while the potential is visible, the overall execution does not achieve the complexity and resilience required for a Level 4 performance.

Across the AI models employed, variations in evaluative strictness were observed, with Qwen consistently producing more conservative assessments. Such differences may be noted in prior research which highlights how differences system design can shape model behavior and judgment (Bender et al., 2021; Liang et al., 2022). Some models appear to apply higher

thresholds when identifying higher-order competencies, requiring more explicit evidence of sustained problem-solving and adaptive processes, while others allow partial or inferred indicators to be recognized. Consequently, the observed differences in AQRF alignment are likely reflective of model-specific evaluative tendencies rather than substantive disagreement about the text itself.

Variations in evaluative outcomes across large language models have been widely noted in prior research, which highlights how differences in training data, alignment objectives, and system design can shape model behavior and judgment (Bender et al., 2021; Liang et al., 2022). Accordingly, differences in evaluative strictness observed in this study are likely attributable to model-specific tendencies rather than substantive disagreement about the text itself.

The engagement of multiple AI chatbots in evaluating the Grade 12 Kurikulum Merdeka English textbook can be conceptualized as a symbolic “intellectual war”, where each system, ChatGPT, Gemini, and Qwen, offers distinct interpretive strengths but collectively demonstrate how the integration of Reading for Emotions (Lian, 2017) and Aesthetic Principles (Ramachandran & Hirstein, 1999) advances textbook evaluation beyond literal comprehension. These frameworks reveal that narratives in textbooks are not simply vehicles of information but also serve deeper functions: they guide emotional resonance, stimulate cognitive engagement, and create aesthetic experiences that make meaning memorable and impactful.

In this sense, writing a textbook becomes both an artistic and cognitive act, where emotional tension, symbolic contrast, and narrative rhythm work together to cultivate reflective thinking rather than rote recall. Learners at the Grade 12 level (typically aged 17–18 years) are expected to possess critical thinking and adaptive reasoning skills in line with AQRF Level 4, which stresses theoretical understanding, independent judgment, and problem-solving in unfamiliar contexts. This expectation resonates with Piaget’s (1969) formal operational stage, where adolescents develop the ability to think abstractly and systematically. Ideally, the Kurikulum Merdeka textbook should already embody this standard, and its implementation can be further strengthened through the BIMA framework (Bridging Intelligence, Mindfulness and Awareness) proposed by Wirawan (2025), which operationalizes multi-dimensional awareness as a foundation for advanced learning. Through BIMA, learning is directed toward knowledge consciousness, where students are not only absorbing information but also cultivating awareness of how knowledge connects to their personal, social, and cultural realities. Such an approach ensures that textbooks serve as bridges between theoretical concepts and real-life application, enabling learners to practice empathy, ethical reasoning, and reflective judgment alongside analytical skills precisely the capacities expected at the formal operational stage.

According to Zull (2002), learning follows a biological cycle in the brain that moves from sensory experience to reflection and action. Moreover, as Zull (2011) explains, learning runs parallel with the fundamental processes by which nervous systems transform sensory experiences into meaningful understanding. Neuroscience research shows that learning emerges through the interaction of sensory experience, emotion, and reflection within neural circuits. When learners engage with emotionally resonant texts, neural pathways associated with perception and reasoning become more active, strengthening comprehension and meaning-making processes. The brain does not merely store information but actively reorganizes experience into meaningful patterns. In line with this perspective, Lotto (2017) argues that perception is not a passive reception of reality but an active construction shaped by prior experience and neural interpretation. The brain does not merely store facts; it continuously forms, manipulates, and reconstructs mental images that represent experience. These images serve as vehicles for comprehension, thought, and action, shaping how learners perceive, feel,

and reason. Thus, every word, sentence, or narrative structure in a textbook becomes a potential neural image, activating patterns of emotion (feeling), interpretation (thinking), and sensory memory (perception). When these three dimensions operate together, learning transcends memorization and becomes an act of meaning-making a process in which knowledge is internalized through emotional resonance and cognitive imagery. In this sense, Freire's (1970) notion of *generative themes* provides the social and existential grounding for such neural processes, as meaningful learning must arise from concrete human realities that stimulate critical reflection. Complementing this, Wirawan's (2025) concept of Regenerative Learning within the BIMA framework extends Zull's neurobiological insight by emphasizing that learning should not only generate meaning but also regenerate awareness, empathy, and moral consciousness through reflective and mindful engagement with lived experience.

By mapping these findings against AQRF Level 4, which emphasizes theoretical understanding, adaptive reasoning, and problem-solving in unfamiliar contexts, AI provides evidence that aligning pedagogy with both emotional and aesthetic cognition is essential for Indonesia's educational transformation. Situated within the broader vision of Indonesia Emas 2045, this process demonstrates that embedding emotional, aesthetic, and cognitive rigor into textbooks is not an optional enhancement but a strategic investment in preparing a generation capable of meeting international standards while staying rooted in a culturally and emotionally intelligent pedagogy.

Conclusion

This paper places AI models, namely ChatGPT, Qwen, and Gemini, at the forefront of new analytical tools in assessing Grade 12 Kurikulum Merdeka English texts, particularly that of Gurung Mali's *The Breakers of Limitations*. Combining Reading for Emotions (Lian, 2017) and neuroaesthetic principles (Ramachandran & Hirstein, 1999), the analysis shows how textual elements set off feeling, thinking, and perception and stresses the interconnected nature of learning: cognitive, emotional, and aesthetic (Zull, 2011). Though the story offers cultural richness, appealing narratives, and partial attainment of AQRF Level 4 competencies, there is limited evidence of higher-order reasoning, reflective awareness, and depth of emotion.

Based on the insights of cognitive development by Piaget (1969) and Zull (2011), this paper emphasizes that learning progresses from concrete experience to abstract understanding through constant reflection. By the formal operation stage, the learner is supposed to achieve *conscientização*, which, according to Freire (1970), is a critical awareness that allows people to perceive themselves as subjects with the capability to change their world. Therefore, English textbooks should inculcate generative themes, that is, themes taken from the real-life contexts of learners to raise consciousness, foster freedom, and promote social responsibility.

AI-assisted evaluation moves, therefore, beyond the realm of technical analysis to a reflective and ethical examination of how education shapes cognition and awareness. Such AI-based evaluations should be extended in future research to cover numerous types of texts, grade levels, and disciplines, including classroom observations and learner outcomes. Future studies are also recommended to include expert human judgement, particularly from practicing teachers, lectures, or curriculum practitioners to interpret the analysis with professional pedagogical insight. Embedding emotional, cognitive, and aesthetic depth into the curriculum will foster critical thinking, reflective judgment, and emotional intelligence some of the competencies needed to realize Indonesia Emas 2045, where education becomes a path toward intellectual and moral liberation. This approach aligns with Sustainable Development Goal 4 and reiterates

a commitment toward ensuring inclusive, equitable, and quality education to enable all learners to thrive as conscious and compassionate global citizens.

Acknowledgement

First and foremost, I would like to express my most sincere gratitude to Lala Bumela, Ph.D., Director of the International Office and Partnership at UIN Siber Syekh Nurjati Cirebon (UIN SSC), for guidance, mentorship, and support throughout the completion of this study. His knowledge of international academic cooperation and commitment to the excellence of education immensely helped with the research direction and methodology. Lala Bumela's facilitation in terms of institutional resources, encouragement to navigate the challenges of the research, and advice on aligning the study with both scholarly standards and practical relevance considerably contributed to the successful completion of this work. His commitment to furthering education and international cooperation has inspired me and built a base for high-quality and impactful research.

In addition, the author would like to extend profound gratitude to the co-authors for their exceptional collaboration and support throughout the research process. Their expertise, critical feedback, and attention to detail enriched the study, especially in integrating AI-assisted evaluation, Reading for Emotions, and neuroaesthetic frameworks within the analysis of the Grade 12 Kurikulum Merdeka English textbook. From conceptualization to drafting, they strengthened methodological rigor, enhanced clarity, and inspired reflective discussions on curriculum design, higher-order thinking, and learner engagement. Their unwavering encouragement and intellectual contribution were invaluable in helping the research not only address theoretical frameworks but yield actionable insights into ways of educational improvement and learner development.

References

- Ansori, R. W. (2021). Ancangan Indonesia Emas 2045 dalam Buku Teks Bahasa Indonesia Terbitan Kementerian, Pendidikan, dan Kebudayaan Tahun 2017. *Jurnal Pendidikan: Riset dan Konseptual*, 5(4), 551–557. https://doi.org/10.28926/riset_konseptual.v5i4.410.
- ASEAN Secretariat. (2012). *ASEAN Qualifications Reference Framework (AQRF)*. <https://asean.org/>.
- Atchley, P., Pannell, H., Wofford, K., Hopkins, M., & Atchley, R. A. (2024). Human and AI collaboration in the higher education environment: Opportunities and concerns. *Cognitive Research: Principles and Implications*, 9(1), 20. <https://doi.org/10.1186/s41235-024-00547-9>
- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? In Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (pp. 610–623). *Association for Computing Machinery*. <https://doi.org/10.1145/3442188.3445922>.
- Bengtsson, M. (2016). How to Plan and Perform a Qualitative Study Using Content Analysis. *Nursing Plus Open*, 2, 8-14. <https://doi.org/10.1016/j.npls.2016.01.001>.
- Cadwell, D., & White, P. R. R. (2017). That's not a narrative; this is a narrative: NAPLAN and pedagogies of storytelling. *Australian Journal of Language and Literacy*, 40(1), 16–27.

- <https://doi.org/10.1007/BF03651981>.
- Carvajal, A. L. P., & Vicente, R. S. (2024). The development and analysis of a framework for enhancing the Philippines' competitiveness in ASEAN surveying. *International Journal of Open-access, Interdisciplinary and New Educational Discoveries of ETCOR Educational Research Center (iJOINED ETCOR)*, 3(1), 401-413. [https://etcor.org/storage/iJOINED/Vol.%20III\(1\),%20401-413.pdf](https://etcor.org/storage/iJOINED/Vol.%20III(1),%20401-413.pdf).
- Celik, I., Dindar, M., Muukkonen, H., & Järvelä, S. (2022). The promises and challenges of artificial intelligence for teachers: A systematic review of research. *TechTrends*, 66(4), 616-630. [The Promises and Challenges of Artificial Intelligence for Teachers: a Systematic Review of Research | TechTrends | Springer Nature Link](#)
- Damayanti, I., & Mukarto, F. X. (2024). The analysis of tasks presented in Kurikulum Merdeka textbook English for Nusantara. *Premise: Journal of English Education*, 13(1), 21. <http://dx.doi.org/10.24127/pj.v13i1.9065>.
- Dewey, J. (1938). *Experience and Education*. New York, NY: Macmillan. <https://archive.org/details/ExperienceAndEducation>.
- EF Education First. (2024). *EF English Proficiency Index 2024*. <https://www.ef.com/assetscdn/WIBIwq6RdJvcD9bc8RMd/cefcom-epi-site/reports/2024/ef-epi-2024-english.pdf>
- Freire, P. (1970). *Pedagogy of the oppressed*. New York, NY: Continuum Publishing. <https://fsi-ebcao.princeton.edu/sites/g/files/toruqfl411/files/media/freire>.
- Gasani, L., Nugraha, M. A., & Purnama, D. W. (2025). Exploring Factors Enhancing English Speaking Proficiency. *Biormatika: Jurnal Ilmiah Fakultas Keguruan dan Ilmu Pendidikan*, 11(1), 52–62. <https://doi.org/10.35569/biormatika.v11i1.2273>.
- Hardini, S. R., Merdianto, A., Marjenny, Nurhayati, R., Syathroh, I. L., & Dadan. (2022). *Bahasa Inggris: Life today untuk SMA/MA kelas XII*. Pusat Perbukuan Badan Standar, Kurikulum, dan Asesmen Pendidikan. <https://buku.kemendikdasmen.go.id/katalog/bahasa-inggris-life-today-untuk-smama-kelas-xii>
- Ji, H., Han, I., & Ko, Y. (2023). A systematic review of conversational AI in language education: Focusing on the collaboration with human teachers. *Journal of Research on Technology in Education*, 55(1), 48-63. <https://doi.org/10.1080/15391523.2022.2142873>.
- Karademir, K., & Alper, A. (2024). Evaluating the evaluators: A comparative study of AI and teacher assessments in higher education. *Digital Education Review*, (45), 124-140. <https://doi.org/10.1344/der.2024.45.124-140>.
- Komite Nasional Kualifikasi Indonesia (KNKI). (2020). *ASEAN Qualifications Reference Framework: Referencing Report of Indonesia* (Jakarta: Kementerian Pendidikan dan Kebudayaan). <http://kkni.ristekdikti.go.id>.
- Krakowski, S. (2025). Human-AI agency in the age of generative AI. *Information and Organization*, 35(1), 100560. <https://doi.org/10.1016/j.infoandorg.2025.100560>.
- Kuroda, K., Sugimura, M., Kitamura, Y., & Asada, S. (2018). Internationalization of higher education and student mobility in Japan and Asia. [Internationalization of higher Education and student mobility in Japan and Asia; 2018](#).

- Lian, A. B. (2017). Reading for emotion with ICT tools. In Chen, W. et al. (Eds.), *Proceedings of the 25th International Conference on Computers in Education*. <https://library.apsce.net/index.php/ICCE/article/view/2191>.
- Lian, A., & Sussex, R. (2018). Toward a Critical Epistemology for Learning Languages and Cultures in Twenty-First Century Asia. In *Multilingual Education (Vol. 24)*. https://doi.org/10.1007/978-3-319-69995-0_3.
- Liang, P., Bommasani, R., Lee, T., Tsipras, D., Soylu, D., Yasunaga, M., Zhang, Y., Narayanan, D., Wu, Y., Kumar, A., Newman, B., Selvaraj, S., Abeysinghe, A., Bernal, E., Mandal, P., Hwang, C., Holtzman, A., Jiang, S., Kalluri, P., ... Stanford CRFM. (2022). Holistic evaluation of language models (arXiv:2211.09110). arXiv. <https://arxiv.org/abs/2211.09110>.
- Lotto, B. (2017). *Deviate: The Science of Seeing Differently* (O. P. Group, Ed.). https://books.google.co.id/books/about/Deviate.html?id=OtjYjgEACAAJ&redir_esc=y.
- Mahligawati, F., Allanas, E., Butarbutar, M. H., & Nordin, N. A. N. (2023, September). Artificial intelligence in physics education: A comprehensive literature review. In *Journal of Physics: Conference Series* (Vol. 2596, No. 1, p. 012080). IOP Publishing. [Artificial intelligence in Physics Education: a comprehensive literature review - IOPscience](https://doi.org/10.1088/1742-6596/2596/1/012080)
- Marliana, R. (2024). Analyzing English's Impact on Global Competitiveness and Sustainable Progress. *Jurnal Riset Ilmu Pendidikan*, 4(4), 255-260. <https://doi.org/10.56495/jrip.v4i4.932>.
- Mouta, A., Pinto-Llorente, A. M., & Torrecilla-Sánchez, E. M. (2025). “Where is agency moving to?”: Exploring the interplay between AI technologies in education and human agency. *Digital Society*, 4(2), 49. <https://doi.org/10.1007/s44206-025-00203-9>.
- Nguyen, A., Gardner, L., & Sheridan, D. (2020). Data analytics in higher education: An integrated view. *Journal of Information Systems Education*, 31(1), 61–71. <http://jise.org/Volume31/n1/JISEv31n1p61.html>.
- Pham, N. N. H. (2025). The Use of ChatGPT in EFL students as a Learning Assistant in their Writing Skills: A literature review. *International Journal of AI in Language Education*, 2(1), 38-54. <https://doi.org/10.54855/ijaile.25213>.
- Piaget, J., & Inhelder, B. (1969). *The psychology of the child*. Basic Books. <https://www.basicbooks.com/titles/jean-piaget/the-psychology-of-the-child/9780465095001/>.
- Ramachandran, V. S., & Hirstein, W. (1999). The Science of Art: A Neurological Theory of Aesthetic Experience. *Journal of Consciousness Studies*, 6, 15-51. <https://philarchive.org/rec/RAMTSO-5>.
- Republic of Indonesia. (2025). Presidential Instruction No. 1 of 2025 on Budget Efficiency. <https://peraturan.bpk.go.id/Details/313401/inpres-no-1-tahun-2025>.
- Russell, S., & Norvig, P. (2021). *Artificial Intelligence: A modern approach* (4th ed.). Pearson. <https://aima.cs.berkeley.edu/>.
- Selwyn, N. (2019). *Should robots replace teachers? AI and the Future of Education*. (1st ed.) Polity Press. <https://research.monash.edu/en/publications/should-robots-replace-teachers-ai-and-the-future-of-education>.

- Senowarsito, Senowarsito & Suwandi, Suwandi & Musarokah, Siti & Ardini, Sukma. (2023). The Analysis of Feasibility in EFL Textbook: A Contribution to Kurikulum Merdeka. Arab World English Journal. 305-319. <https://dx.doi.org/10.24093/awej/call9.21>.
- Sudimantara, L. B., Baehaqi, L., Lian, A., & Lian, A. (2025). Multisensory CALL for Under-Resourced Universities and Schools in Indonesia. Palgrave Macmillan. <https://link.springer.com/book/10.1007/978-3-031-91387-7>.
- Sudarma, U. (2022). Pendidikan karakter dalam mewujudkan sumber daya manusia berdaya saing menuju Indonesia Emas 2045. *Sharia: Jurnal Kajian Islam*, 1(1), 37–55. <https://doi.org/10.59757/sharia.v1i1.4>.
- Tenny, S., Brannan, J., & Brannan, G. (2025). Qualitative Study. *StatPearls*. <http://www.ncbi.nlm.nih.gov/books/NBK470395/>.
- Trikoili, A., Georgiou, D., Pappa, C. I., & Pittich, D. (2025). Critical Thinking Assessment in Higher Education: A Mixed-Methods Comparative Analysis of AI and Human Evaluator. *International Journal of Human-Computer Interaction*, 1–14. <https://doi.org/10.1080/10447318.2025.2499164>.
- Urmeneta, A., & Romero, M. (2025). AI as a creative partner: A PRISMA review of AI's role in supporting creativity in education. In *Frontiers in Education*, 10, 1602151. <https://doi.org/10.3389/feduc.2025.1602151>.
- Wachyuniah, Y., Wirza, Y., & Karnedi, K. (2024). Comparison Of Two English Textbooks Used in Indonesia Secondary Education: English as a Second Language (Cambridge Igcse) Vs English for Nusantara (Kurikulum Merdeka). *Klasikal: Journal Of Education, Language Teaching And Science*, 6(3), 957-968. <https://doi.org/10.52208/klasikal.v6i3.1244>.
- Williamson, B. and Eynon, R. (2020) Historical Threads, Missing Links, and Future Directions in AI in Education. *Learning, Media and Technology*, 45, 223-235. <https://doi.org/10.1080/17439884.2020.1798995>.
- Wirawan, R. (2025). *From Seci to Bima – Toward regenerative knowledge management through bridging intelligence, mindfulness, and awareness* [White paper]. International Society for the Systems Sciences Conference.
- Zull, J. E. (2011). *From brain to mind: Using neuroscience to guide change in education* (1st ed.). Stylus Publishing. <https://doi.org/10.4324/9781003444886>.
- Zull, J. E. (2002). *The art of changing the brain: Enriching the practice of teaching by exploring the biology of learning* (1st ed.). Stylus Publishing. <https://doi.org/10.4324/9781003447573>.

Biodata

Author 1 – Muhammad Azkiya Bahtsulkhoir

I am Muhammad Azkiya Bahtsulkhoir, 20, a 5th-semester pre-service English teacher at UIN Siber Syekh Nurjati Cirebon, Indonesia. I am passionate about language teaching, digital content creation, and educational innovation. My research interests include integrating technology in EFL learning, creative storytelling, and enhancing student engagement through innovative pedagogical approaches.

Co-Author 1 – Salsabilla

I am Salsabilla, 20, a 5th-semester pre-service English teacher at UIN Siber Syekh Nurjati Cirebon, Indonesia. I am interested in educational technology and curriculum development. My research focuses on applying digital tools to support student-centered learning and fostering interactive, meaningful EFL experiences for learners in Indonesian classrooms.

Co-Author 2 – Ananditamala Syalwa Heryana

I am Ananditamala Syalwa Heryana, nearly 20, a 3rd-semester pre-service English teacher at UIN Siber Syekh Nurjati Cirebon, Indonesia. I am passionate about language learning and educational technology. My research interests include innovative teaching strategies, student-centered approaches, and developing meaningful EFL learning experiences to advance education in Indonesia.