Vol. 15, No. 04, pp. 2698-2711, 2025 DOI: 10.23960/jpp.v15i4.pp2698-2711

Jurnal Pendidikan Progresif

e-ISSN: 2550-1313 | p-ISSN: 2087-9849

http://jpp.fkip.unila.ac.id/index.php/jpp/

Enhancing EFL Reading Comprehension and Engagement through Multimodal Instructional Design: A Convergent Mixed-Methods Study

Rismar Riansih[1,*], Safnil Arsyad[1], & Zifirdaus Adnan[2]

[1]Doctoral Program of Applied Lingistics, Universitas Bengkulu, Indonesia

[2]Indonesian Studies, School of Humanities, Arts & Social Sciences (HASS), University of New England, NSW, Australia

*Corresponding email: wonder.lady28@gmail.com

Received: 21 November 2025 Accepted: 20 December 2025 Published: 25 December 2025

Abstract: Enhancing EFL Reading Comprehension and Engagement through Multimodal Instructional Design: A Convergent Mixed-Methods Study. **Objectives:** This study explores to what extent the multimodal features incorporated into Indonesian high school English textbooks impact students' engagement and comprehension during reading. As students engage more and more with visually oriented, digitally mediated texts, it becomes imperative to understand how multimodal materials can shape learning. **Method:** "A convergent parallel mixed-method research design was employed. A total of 75 government-issued English course Grade XII students participated in the study. The researchers collected quantitative data through pre-tests, post-tests, and a student engagement questionnaire that measures behavioral, emotional, and cognitive engagement. Qualitative data were collected from an open-ended questionnaire that investigated how participants interpreted the textbooks' images, colors, layout, and other multimodal resources. **Findings:** The results indicated that students' comprehension was significantly improved by multimodal texts, as evidenced by their mean scores compared to traditional texts-only. The students noted that graphic aids assisted their comprehension of difficult language by reinforcing meaning and information recall. The conjunction of multimodal features enhanced motivation and reduced boredom, leading to greater emotional engagement. The qualitative answers also suggested that students found the multimodal layouts more aesthetically appealing, more organized and easier to move along with during a positive reading experience. In summary, the combination of quantitative and qualitative findings shows that multimodal features do significantly contribute to improving students' engagement and understanding. **Conclusion:** The findings suggest that English teachers in Indonesia should incorporate a more systematic use of multimodal design in their integration of visual and spatial aspects and help their students question images critically. Multimodal literacy should be embedded in teachers' pedagogy to create education that is more relevant, inclusive, and effective in today's EFL classrooms.

Keywords: multimodal texts, english textbooks, student engagement, reading comprehension, EFL.

To cite this article:

Riansih, R., Arsyad, S., & Adnan, Z. (2025). Enhancing EFL Reading Comprehension and Engagement through Multimodal Instructional Design: A Convergent Mixed-Methods Study. *Jurnal Pendidikan Progresif*, *15*(4), 2698-2711. doi: 10.23960/jpp.v15i4.pp2698-2711.

INTRODUCTION

Research on multimodal learning has shown that making sense occurs not just through linguistic text but also through the interplay of visual images, spatiality, and design, working together to support understanding. However, when the world of multimodal literacy is in high demand globally, Indonesian English textbooks are surprisingly

stodgy and classical on paper, words-heavy tomes packed with long blocks of text and laid thinly on a page. This has also been shown by some works, which highlight that students find such resources helpful in addressing the challenge of understanding and manipulating, particularly around abstract concepts or new phrases. Poor visual support increases cognitive load during learning, as learners are unable to form mental images of what they are learning.

According to these worries, an imbalance emerges between the multimodal worlds children engage in outside of school and the monomodal worlds teachers provide in their classrooms. According to this developing position in educational research, multimodality goes beyond traditional textual (print) modes and includes digital and interactive means of meaning-making. Therefore, the view of discourse, viewing, and the image (as grammar) generated and directed by online screen communication shifted to encompass dynamic multimodal arrays of 3D images with user brainwave activity that served real-world originality, once tied to current communication behavior of viewing through networked digital world space. For this reason, too, readers are in this day and age expected to interpret hyperlinks, animations, digital design, and multimodal interfaces in new ways that require new literacies.

The new philosophies that were offered by recent models, such as digital multiliteracies (Cope & Kalantzis, 2020), multimedia education in online learning environments (Rowsell & Walsh, 2011), and Mayer's reconceptualization of Multimedia Learning Theory (2001) highlight that teaching quality today depends on learners interacting with these semiotic resources to both understand and complement meaning across verbal, visual, spatial, and digital semiotic resources. These fresh insights assert that multimodal literacy is about more than the decoding of static images, but is an understanding of how digital modes are combined and re-mediated to foster learning in digitally mediated settings.

For teaching and learning in this multimodal environment to be possible, literacy has been conceptualized as the capacity to read between lines as well as between spaces beyond letters (Bezemer & Jewitt, 2010; Unsworth & Ngo, 2021). Multimodal pedagogies are now increasingly acknowledged for aiding understanding, for example, by negotiating multimodal language, visual, and spatial modes simultaneously (Cope & Kalantzis, 2009; Tseronis, 2018). There is also a suggestion that students process visual-verbal combinations more successfully than they handle linguistic information in an isolated manner (Dalton & Proctor, 2007; Danielsson & Selander, 2016a), and it is evident that textbooks should be designed with the capability to align meaningful multimodal capabilities.

With this consideration, the concept of multimodal literacy has become an area that might be pivotal within the field of English language teaching (ELT). Jewitt (2009) maintains that students today should be literate, as well as understanding, in terms of the language code and visual elements in communication, as well as in multimodal decoding of signs; they should be familiar with the current academic texts as they compare to the written world and in normal communication. Serafini (2012) further argues that multimodal texts afford more affordances of comprehension for visual learners and students with various learning needs.

Alternatively, in other words Multi-literacy textbooks can be sensitive to how students learn in today's media-rich environments. These theoretical forecasts find empirical support. As responses in the current study indicated, EFL students exposed to multimodal reading texts for learning English showed superior performance in terms of critical thinking skills compared to

monomodal learners' performance and achievement when exposed (Yeh & Lai, 2019). Ajayi (2012) also found that there seemed to be a higher cognitive and emotional involvement compared to multimodal media by students when they had multimodal and multitampled use of the two texts. In the case of the Asian context, the pictorial support and layout of English textbooks may have helped to stimulate students' motivation and comprehension (Hamidah, 2022; Huda et al., 2025; Lim et al., 2022; Muhassin et al., 2022).

However, we encountered scattered, descriptive research on multimodality in Indonesia. Although international research on multimodality in particular demonstrates how multimodal resources promote the development of meaning by media materials (Bezemer & Kress, 2016; Jewitt, 2014; Kress & Van Leeuwen, 2006), previous studies in the Indonesian context have shown that literary texts are mainly studied for the textual features only (grammatical sequencing), textuality structure, and vocabulary load, not for the possibility of taking multimodal elements in the text into consideration (A. Andita & Pipit, 2024; Emilia & Hamied, 2019). The multimodal materials in English government textbooks are rarely reported even more generally by those who examine the visual elements, and who rather merely accept them superficially (Elmiana, 2019; Jamilah & Mutia Ismail, n.d.). Evidence in studies from Southeast Asian EFL classrooms suggests that multimodal elements are not utilized systematically or used for a meaningful purpose (O'Halloran, 2004).

Research has shown that many of the visual, spatial, and typographical components in EFL materials are aesthetic decorations or barely aligned with teaching and learning objectives. For instance, badly-designed multimodal resources restrict the students' sense-making activities, while a well–designed multimodal material improved students' comprehension and engagement (Kaowiwattanakul, 2025). These findings about the multimodal aspects of Indonesian English textbooks contribute to an anxious perception that they have not been formed yet, in order to unlock their potential for promoting literacy learning with greater depth and quality (Unsworth & Ngo, 2021).

According to Fitriana & Wirza (2021), multimodal features that appear in any secondary school-level textbooks are poorly integrated and do not help students approach or comprehend the content. They do not effectively leverage multimodal content, so students cannot easily engage with the text. Worse, this body of work has not yet explored the impact of multimodal features on students' engagement or comprehension in real classrooms (Jewitt, 2014), leaving a significant empirical gap in the practice of multimodal pedagogy in Indonesia. Thus, until this phase, the literature on the impact of different multimodal resources on student engagement and understanding of reading in a classroom setting in Indonesia is limited.

In the absence of this analysis, this gap underscores the need for comparative analyses not just of multimodal factors, but also of actual learning outcomes, particularly for textbooks with extensive impact across the whole nation. However, only a few works have addressed multimodality in an Indonesian context. Notwithstanding rich evidence on the effectiveness of instructional materials in language education globally (Richards, 2006) and the demonstrated value of multimodal materials for learning (Kress & Jewitt, 2010). Some recent studies have identified a tendency toward linguistic claustrophobia in Indonesian textbook-based research, with a stronger focus on linguistic features rather than multimodality in language pedagogy (e.g., grammar sequencing, text mapping, and form load) (Fitriana & Wirza, 2021). Multimodal resources in government

written English textbooks are found in limited literature and, if present in those textbooks, little pedagogical aspect of the phenomenon or its impact on student learning is mentioned. Other textbook multimodal studies have reported similar profiles. Quang Dong et al. (2024) found that visuals were decorative or circumstantially connected to the surrounding verbal text in a medical textbook and serve as diminished semiotic resources in making meaning. Similar discoveries are also seen in the literature of other countries (like Indonesia). For example, Unsworth & Mills (2020) observed in the secondary school textbooks that images could not serve a cognitive function but rather an ideological or illustrative one to students' learning process, such that, when images are not designed and positioned pedagogically, they do not further learning. In general, Osborn (2017) discovered visual resources in world history textbooks that were biased towards the support of authoritative narratives when not understood, and recommended that multimodal sources may lack their pedagogical potential unless they were explicitly created for such ends. Building on this argument, Kress & Jewitt (2010) proposed multimodal texts that "work" recognise that the strategic utilisation of modes – image, layout, and typography – is designed to help learners to make meaning, and yet principled multimodal design has been underexploited in many school texts. In summary, these studies suggest that while multimodal resources exist in textbooks, there is no empirical work exploring the impact of multimodal features on both students' engagement and understanding during instruction. This gap speaks to the need for comparative studies of the multimodal features being analyzed that do not rely solely on their descriptive qualities to examine whether and how they influence learning.

To address this gap, the current research has employed a mixed-methods study to investigate the effects of multimodal features in high school English texts (including those reported or suggested by Kemendikbudristek) on students' engagement with and understanding of the texts they have read. This research aims to develop a holistic understanding by integrating quantitative data from surveys and comprehension tests with qualitative data from focus group discussions to fully understand the impact of efforts to integrate multimodal texts into students' reading. Another advantage of capturing students' voices in the data through focus groups is the insight into how students render multimodal texts in their own language. The strength of the present study lies not only in using authentic textbook materials but also in blending different analytical viewpoints to examine the multimodal dimensions of government-issue English textbooks. While previous studies tended to contrast textual explanation with an isolated learning outcome that analyzes the interaction between students' engagement and comprehension, this study integrates quantitative assessment of how effective multimodal practices are with learners' qualitative responses, providing a more holistic picture of what multimodality does in practice within the classroom.

Methodologically, the convergent mixed-methods design allows the convergence of measurable performance data with learners' interpreting experience in this research, an approach rarely employed in multimodality research with Indonesian students. On an analytical level, the present paper engages with multimodal elements not only as decorations or structures of a text, but as meaning-making resources enabling cognitive and affective engagement. Theoretically, it contributes to the existing multimodal literacy debates by focusing on MM learning in the Indonesian secondary school socio-educational context, which is based on very prescribed texts. Collectively, these contributions advance a distinct, contextually situated development in multimodal literacy studies within the Indonesian context. Based on these

considerations, the study is guided by the following research questions:

1. To what extent do multimodal elements in government-issued English textbooks affect Indonesian high school students' engagement and reading comprehension?
2. How do students perceive and respond to the multimodal features present in these English textbooks?

By answering these questions, this study aims to inform future textbook development, teacher training, and curriculum planning in Indonesia, especially as the country moves toward a more digitized and competence-based educational model under the Merdeka Belajar policy.

METHOD

Participants

This study involved 75 Grade XII students enrolled in an English course at a public senior high school in Indonesia. All participants completed both the pretest and posttest phases and were therefore included in the final analysis. The participants were selected using a total sampling technique, as all students in the class met the study's inclusion criteria. While convenience sampling limits the generalizability of the findings, it is a common methodological design of school-based research that relies on administrative consent and classroom availability to recruit participants (Cohen et al., 2018). The participants exhibited varying levels of English proficiency, and thus, different viewpoints were reflected in the quantitative measures and responses. After the quantitative experiment, 12 students were invited to participate in the qualitative study.

Research Design and Procedure

This research used a convergent parallel mixed-methods design where quantitative and qualitative data were collected and analyzed concurrently and then integrated to provide a holistic perception of the research issue (Creswell, 2014; Creswell & Clark, 2018). This model was chosen to describe and measure both observable trends in students' engagement and reading comprehension, and to provide detailed, multifaceted descriptions of students' responses to multimodal English textbook materials. Mixed methods research is particularly suitable for educational applications, which make learning complex and multi-dimensional and cannot be entirely accounted for by a single methodological approach.

Formally seeking permission before conducting the research, the school prepared two versions of the reading materials: a traditional text-only version and a multimodal version with images, color cues, and layout enhancements. All instruments were used through a low-scale trial (n = 10) to ensure clarity and appropriate difficulty. Students completed the Student Engagement Questionnaire for quantitative analysis and also took the traditional reading comprehension test. One week later, the same students were administered the multimodal reading comprehension test to reduce possible recall effects. During the qualitative stage, the students completed an open-ended questionnaire to describe their perceptions of the textbook's multimodal features, and a subsequent interview was scheduled to further analyze and interpret the written responses from the selected participants. Quantitative and qualitative results were analyzed separately and then combined in a comprehensive interpretation of the relationships between multimodal features and students' engagement and comprehension.

Data Collection Instruments and Procedures

Student Engagement Questionnaire (Non-Test Instrument)

The students' engagement was determined by a questionnaire adapted from the Appleton (2006) Student Engagement Instrument (SEI).

The survey comprised 18 items across three dimensions of engagement: behavioural, emotional, and cognitive. Each dimension was measured using six items rated on a five-point Likert scale ranging from strongly disagree (1) to strongly agree (5). Test items were adapted to the context of multimodal reading in English textbook materials while remaining closely aligned with the original theoretical constructs.

Table 1. Indicators and items distribution

Indicator	Description	Number of Items	Example Item
Behavioural Engagement	Students' effort, participation, and persistence while completing reading tasks	6 items	"I try hard to understand the reading passages even when they are difficult."
Emotional Engagement	Students' interest, enjoyment, and affective response toward the reading materials	6 items	"I feel more interested in reading when the text includes pictures or colours."
Cognitive Engagement	Students' investment in understanding concepts and using strategies when reading	6 items	"I use clues from images and layout to help me understand the text."

Validity and Reliability

A pair of educational psychologists and English teachers evaluated the adequacy, relevance, and clarity of the items to determine content validity. Construct validity was enhanced by associating items with the theoretical characteristics of involvement as outlined in the SEI framework (Skinner et al., 2009; Skinner et al., 2009). Cronbach's alpha was calculated to assess internal consistency reliability, which was a value of 0.87, indicating a good degree of reliability. Before the main data collection, the questionnaire was piloted with ten students, and minor modifications were made to improve response consistency and clarity.

Reading Comprehension Test (Cognitive Test Instrument)

A teacher-designed reading comprehension test was created to measure students' comprehension of reading materials presented in two parallel formats: a traditional text-only version and a multimodal version. Both formats used the same verbal text for comparison across conditions. Five multiple-choice questions were completed after each passage, covering a total of ten items that measured similar concepts: naming main ideas, understanding details, making inferences, understanding vocabulary, and interpreting textual or visual support.

The multimodal version was designed by the researcher through the integration of pedagogically grounded visual and layout elements, including images, colour cues, and structured page layouts. The selection and design of these multimodal features were informed by principles of multimedia learning and multimodal literacy. Images were chosen to directly represent key concepts, events, or vocabulary items described in the text, to support inferential comprehension and meaning construction. Colour cues were used selectively to highlight headings, key information, and transitions between ideas, to guide students' attention and reduce cognitive load rather than serving decorative purposes. Layout features, such as image placement adjacent to relevant text segments, spacing, and clear sectional organisation, were designed to support text navigation and coherence, consistent with multimodal design perspectives (Bezemer &

Kress, 2016; Danielsson & Selander, 2016a; Mayer, 2009).

The test items were designed in accordance with the Grade XII curriculum and the English textbooks used in the class. Content validity was determined through review by two English education professionals to ensure alignment with curricular objectives and multimodal learning principles. The instruments were piloted with a small group of students (n = 10) outside the main sample to assess item clarity and difficulty. Reliability analysis using the KR-20 formula yielded a coefficient of 0.82, showing strong internal consistency with this comprehension test.

Table 2. Indicators and item distribution

Indicator	Description	Number of Items per Test	Example Item
Main Idea Identification	Ability to identify the central idea of the passage	1 item	"What is the main topic discussed in the text?"
Detail Comprehension	Ability to recognize important factual information	1 item	"Which detail is mentioned in the second paragraph?"
Inference Making	Ability to draw logical conclusions beyond explicit information	1 item	"What can be inferred from the author's explanation?"
Vocabulary Meaning	Understanding word meaning from context or visual cues	1 item	"The word '…' in the text most likely means…"
Interpretation of Structure/Visual Support	Understanding relationships between text and visuals (for multimodal version)	1 item	"How does the image support the explanation in paragraph 3?"

Total: 5 indicators × 1 item each = 5 items per test.

Reading comprehension test for students on written passages was constructed to evaluate students' understanding of the materials in two formats: traditional text-only passages and multimodal passages with visual and layout features. Following each passage came five multiple-choice questions assessing key comprehension indicators, where students' scores were thus directly compared across the two formats. The two versions of the test were reviewed by two experts in English education for content validity regarding the relevant curricular objectives and reading comprehension constructs. The instruments were subsequently trialed with a small group of students (n = 10) outside of the main sample to assess item clarity, difficulty level, and reliability, with minor revisions to this design made. Reliability analysis was then performed to confirm the internal consistency of the comprehension measures.

Open-Ended Questionnaire (Non-Test Instrument)

A semi-structured questionnaire was constructed to capture rich qualitative data about students' perceptions of multimodality in their English textbooks. The instrument comprised five open-ended questions that asked students to reflect on various aspects of their reading experience, such as the perceived usefulness of images, colours, and layout in supporting comprehension; the extent to which multimodal components influenced their motivation to read; and any difficulties encountered when interacting

with multimodal texts. The questions also prompted students to compare their experiences with reading multimodal texts to those with traditional print-only materials and to express their preferences for different reading formats. The questionnaire was quality-checked through expert review done by two qualitative research specialists and piloted with a small number of students to ascertain clarity, relevance, and comprehensibility (Clark, 2010; O'Cathain & Thomas, 2004). To provide a clearer illustration of the qualitative data collected, examples of the questions included:

(1) *"How do pictures, colours, or layout in your English textbook help or hinder your understanding of the reading text?"* and
(2) *"How do you feel when reading a text that includes images and colours compared to reading a text with only words? Please explain your reasons."*

These questions aimed to elicit reflections on the cognitive and emotional aspects of engagement and on students' strategies for making meaning from multimodal texts. Qualitative data were mainly gathered through an open-ended questionnaire designed to explore how students perceived and responded to the multimodal elements in their English textbooks. While this instrument enabled students to express their views and feelings in written form, the depth and richness of the qualitative discussion were further enhanced through limited follow-up interviews with a selected group of participants. This supplementary step was designed to overcome potential limitations of open-ended questionnaires and to elicit more nuanced descriptions of how layout, colours, and images influenced students' motivation, engagement, and comprehension. The follow-up interviews also helped clarify responses that were not fully articulated in writing, thereby improving the interpretation of the qualitative findings.

Many techniques were implemented to improve the validity and trustworthiness of the qualitative findings. Content validity was initiated first by two qualitative research experts, who assessed the clarity, relevance, and alignment of the questions with the study's objectives. Second, methodological triangulation was applied by comparing data obtained from open-ended questionnaires and follow-up interviews. Third, peer debriefing was conducted with another researcher to examine the coding process and to ensure the consistency and coherence of the emerging themes. These procedures helped ensure that the qualitative interpretations accurately represented students' perspectives.

Qualitative data were analyzed using the interactive model proposed by Miles & Huberman (1994). The analysis consisted of three closely linked phases: data reduction, i.e., repeatedly reading students' written and interview responses, coding them, and condensing them into meaningful units; data display, where codes and categories were organised into thematic matrices to facilitate interpretation; and conclusion drawing and verification, during which emerging themes were refined and validated through constant comparison across data sources. This methodical and iterative process helped ensure that the emergent themes were data-driven and grounded in students' actual experiences.

Data Analysis

Qualitative data derived from the open-ended questionnaires and follow-up interviews were analysed using a thematic analysis, following the interactive model described by Miles & Huberman (1994). The analysis began with data reduction, in which students' written and oral responses were read several times, coded, and condensed into meaningful units representative of their experience with multimodal textbook features. These initial codes formed tentative categories in the preliminary themes concerning

multimodal support, motivation, challenges, and reading preferences.

After the data was reduced, the analysis continued to data presentation: codes and categories were arranged in matrices and thematic charts to facilitate systematic comparison across participants. Such a systematic organization facilitated the recognition of recurring patterns and relationships among themes, leading to clearer interpretations of how students perceived and interacted with multimodal content in their reading materials. The data had a visual hierarchy that enabled cross-case analysis and enhanced the coherence of thematic interpretations.

In addition, conclusions were drawn during the conclusion and verification stages, during which the study's emerging themes were fine-tuned through continuous checks against original interview data, with triangulation with interview notes carried out on several occasions. To establish the credibility of qualitative findings, validation procedures were used, including expert validation by two qualitative research experts, methodological triangulation across data sources, and peer debriefing to verify coding decisions and thematic stability. Moreover, all responses were anonymized to minimize self-censorship and protect participants' identities (Nowell et al., 2017). They ensured the qualitative findings, through these methodical processes, were reliable and grounded in students' subjective views and real-world experiences with the topic.

RESULT AND DISCUSSION

This section presents the results and discussion organised by the research hypotheses. For each hypothesis, quantitative and qualitative findings are reported together to provide an integrated interpretation of how multimodal features influence students' engagement and reading comprehension.

Hypothesis 1: Multimodal texts improve students' reading comprehension compared with traditional text-only materials.

Quantitative Analysis

The quantitative results show that students had better reading comprehension when exposed to multimodal texts than when reading only text. Based on visual inspection of the various performance patterns in Figure 1, it is clear that in the multimodal condition, the majority of students score higher than their posttest scores in the traditional condition. Most points lie above the diagonal line representing the baseline, indicating that individual comprehension performance enhanced by multimodal features.

The findings overall indicate that the inclusion of images, colour cues, and a structured layout in reading materials was significantly associated with students' comprehension outcomes. The effect size suggests that the observed change was not only statistically significant but also practically significant in classroom learning situations.

Qualitative Analysis

Qualitative findings explain how reading multimodal texts led to better comprehension among students. Many participants reported that images helped individuals provide a foundation for making sense of unfamiliar vocabulary and for inferring when the written text was difficult to comprehend. Colour cues and page layout also indicated a positive role for the attention shift in highlighting more important information and organising text.

Students reported that these features helped elucidate the text and make the reading process more manageable by helping them understand the thoughts they were reading, minimising confusion between their thoughts and the reading flow. These results, combined, indicate that multimodal

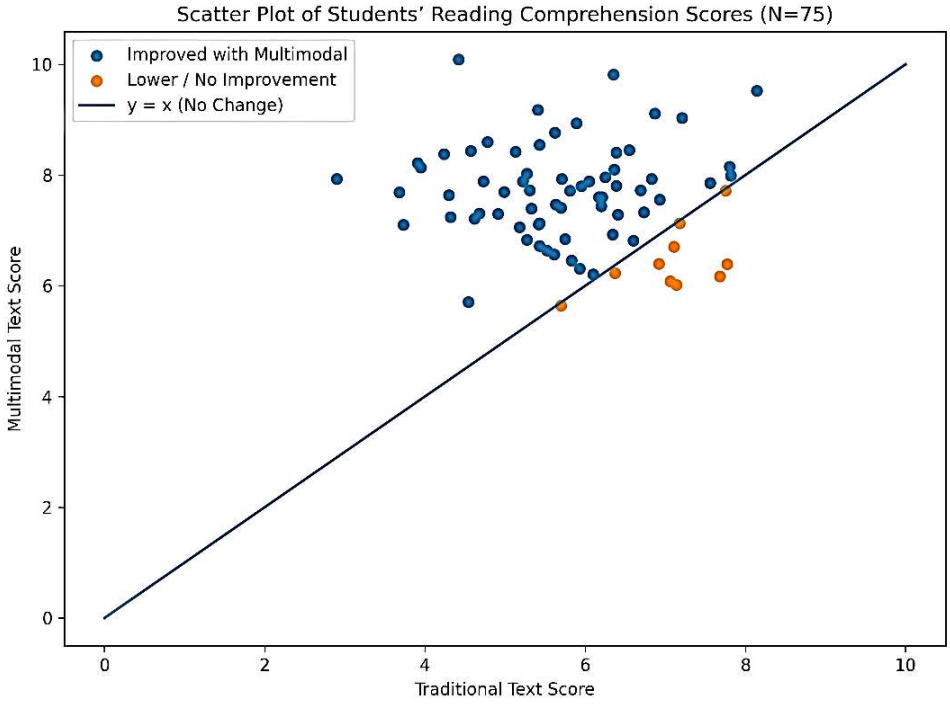

Figure 1. Scatter plot of students' comprehension scores in traditional and multimodal reading conditions (N = 75)

elements served as cognitive scaffolds for meaning formation and for maintaining reading focus, corroborating the quantitative results.

Hypothesis 2: Multimodal Features in Textbooks Positively Influence Students' Engagement in Reading and Learning Activities

Quantitative Analysis

The results of the Student Engagement Questionnaire suggest that across multiple levels of engagement, students' interactions with multimodal English textbook materials led to positive outcomes in behaviour, emotions, and cognition. Figure 2 shows that emotional engagement displayed the highest mean score, followed by behavioural and cognitive engagement.

That this pattern existed indicates that multimodal elements not only motivated students to participate in reading tasks but also heightened their interest and sustained their mental effort. The general quantitative findings indicate a positive relationship between the use of multimodal elements and students' engagement in reading activities.

Qualitative Analysis (Supporting Evidence)

Qualitative data, therefore, provide further insight into the role certain multimodal features played in students' engagement. In contrast to playing a high-level role in engagement, the student cited imagery, colour cues, and layout elements as fulfilling separate but interrelated functions of reading.

Particularly, images were frequently viewed as visuals supports that allowed students to create meaning, especially when faced with unfamiliar vocabulary or complex ideas. Colour cues and layout features encouraged focus by drawing

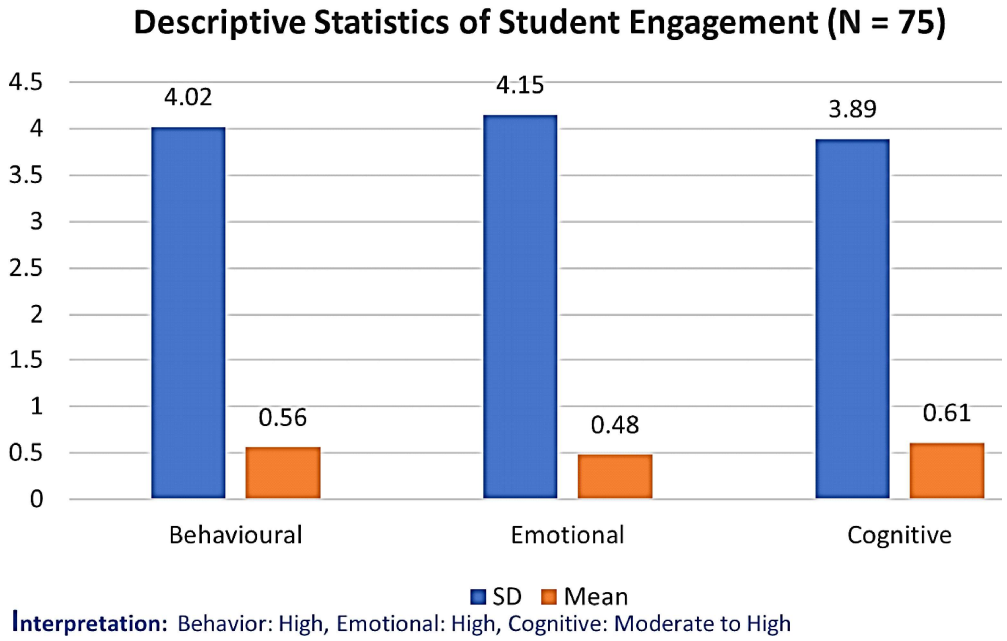

Figure 2. Mean scores of students' behavioural, emotional, and cognitive engagement with error bars representing standard deviations (N = 75)

attention to key topics and by clearly organizing content in the text. These elements also reduced the effort required to locate important points, encouraged students' cognitive involvement, and facilitated active involvement.

In addition, students reported that well-structured and aesthetically appealing layouts made reading less boring and more motivating. These visuals also motivated extended reading effort and emotional engagement. Taken together, the qualitative study shows that multimodal features accounted for the behavioural, cognitive, and emotional dimensions of engagement as expected from these quantitative findings.

The research of this study shows that multimodal text language design in EFL can have a positive effect on students' performance and reading comprehension. The pairing of images, colours, and an organised layout with verbal text enhanced the accessibility, aesthetics, and thoughtfulness of reading tasks. These results align with multimodal literacy views, which argue that meaning-making emerges from a range of semiotic resources rather than solely linguistic texts (Bezemer & Kress, 2016; Jewitt, 2009).

In relation to engagement, the findings align with Fredricks, Blumenfeld, and Paris's (2004) framework, which views engagement as a multilevel construct with behavioral, emotional, and cognitive levels. This study suggested that multimodal features promoted behavioural persistence and cognitive effort (improved organisation and visual guidance) and emotional engagement (increased interest and less boredom).

With respect to reading comprehension, results are in line with Mayer's (2009) Cognitive theory of multimedia learning, which suggests that verbal and visual representations can be integrated to facilitate better understanding. Visual and layout cues were shown to help students infer meaning and understanding, interpret unfamiliar vocabulary, and retain information by minimising cognitive load and supporting visual–verbal understanding (Dalton & Proctor, 2007; Danielsson & Selander, 2016b).

However, there are other reasons to think too. One explanation for the observed improvement might be the novelty effect, which means that the presence of visual content in multimodal texts had a temporary positive effect on students' attention and motivation. Additionally, the efficacy of multimodal support might be affected by learners' learning capacities, as lower-performing readers may depend more on visual and layout cues (Yeh & Lai, 2019).

Most pertinent, this supports prior studies concluding that multimodal aspects alone, by definition, are not optimal and should have intentional instructional design to reinforce instructional aims (Bezemer & Jewitt, 2010; Kress & van Leeuwen, 2006). The present study emphasizes the use of multimodal design embedded in textbooks and methodology. In EFL classroom pedagogy, multimodal design is crucial for optimizing learners' engagement and understanding.

■ CONCLUSION

Certain multimodal elements in the English textbook: images to explain not familiar words, colour codes to present important data, and patterns of page layouts that organize ideas visually, were found to add significantly to engagement and reading comprehension. Students not only achieved a better understanding when reading multimodal texts, but also found these textual features more pleasant and less cognitively demanding to read, as well as easier to maintain attention on the content.

These multimodal features appeared to support comprehension by lowering cognitive load, focusing learners' attention on critical information, and adding visual scaffolding for meaning-making from the text, the researchers discovered. At the same time, emotionally engaging design elements enhanced students' motivation to concentrate on reading assignments. Together, these affordances explain why MM-working was superior to RL-working in supporting comprehension.

The results highlight the necessity of intentional, pedagogically meaningful multimodal design in English textbooks, as well as the strengthening of teacher education in multimodal literacy. Further studies might investigate how digital or interactive multimodal materials enhance comprehension and how multimodality might be applied to other areas of the EFL curriculum.

■ REFERENCES

Ajayi, L. (2012). How teachers deploy multimodal textbooks to enhance English language learning. *TESOL Journal, 6*(1), 16–35.

Andita, P., & Pipit, N. (2024). Investigating the impact of multimodal text on Students' reading comprehension in English learning. *JEELL (Journal of English Education Linguistics and Literature).* https://doi.org/10.32682/cnhfed98

Appleton, J. J. (2006). Measuring cognitive and psychological engagement: Validation of the Student Engagement Instrument. *Journal of School Psychology, 44*(5), 427–445.

Bezemer, J., & Jewitt, C. (2010). Changing textbook designs and learning. *Teaching and Teacher Education, 26*, 1137–1144. https://doi.org/10.1016/j.tate.2009.11.006

Bezemer, J., & Kress, G. (2016). *Multimodality, learning and communication: A social semiotic frame*. Routledge.

Clark, V. L. P. (2010). The adoption and practice of mixed methods: U.S. trends In federally funded health-related research. *Qualitative Inquiry, 16*(6), 428–440. https://doi.org/10.1177/1077800410364609

Cohen, L., Manion, L., & Morrison, K. (2018). *Research methods in education* (8th ed.). Routledge.

Cope, B., & Kalantzis, M. (2009). Multiliteracies: New literacies, new learning. *Pedagogies: An International Journal, 4*(3), 164–195. https://doi.org/10.1080/15544800903076044

Cope, B., & Kalantzis, M. (2020). *Making sense: Reference, agency, and structure In a grammar of multimodal meaning*. Cambridge University Press. https://doi.org/10.1017/9781108882391

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). SAGE Publications.

Creswell, J. W., & Clark, V. L. P. (2018). *Designing and conducting mixed methods research* (3rd ed.). SAGE Publications.

Dalton, B., & Proctor, C. P. (2007). Reading as situated language: A multimodal perspective. *Journal of Adolescent & Adult Literacy, 50*(4), 310–318.

Danielsson, K., & Selander, S. (2016). Reading multimodal texts for learning: A model for cultivating multimodal literacy. *Designs for Learning, 8*(1), 25–36. https://doi.org/10.16993/dfl.72

Elmiana, D. S. (2019). Pedagogical representation of visual images in EFL Textbooks: A multimodal perspective. *Pedagogy, Culture & Society, 27*.

Emilia, E., & Hamied, F. A. (2019). Systemic functional linguistic genre pedagogy in Indonesian schools. *TEFLIN Journal, 26*(2), 155–180.

Fitriana, W., & Wirza, Y. (2021). An analysis of multimodal text in EFL textbook of secondary school in Indonesia in assisting students' text understanding. *In Proceedings of the Thirteenth Conference on Applied Linguistics (CONAPLIN 2020)*. Atlantis Press.

Hamidah, M. (2022). A multimodal analysis of English textbook for vocational high school: A study of English for specific purposes. *Proceedings of the English Language Teaching, Literature, and Translation (ELTLT), 11*(1). https://proceeding.unnes.ac.id/index.php/ELTLT/article/view/162

Huda, T., Loan, N. T. T., Prismantikasari, A., Rachmanda, R., Aroeboesman, F. P., & Adhinata, A. A. (2025). Multimodal gender representation in Indonesian EFL textbooks. *JEES (Journal of English Educators Society), 10*(1). https://doi.org/10.21070/jees.v10i1.1926

Jamilah, S., & Ismail, N. M. (2024). Multimodal analysis of an English textbook used for EFL young learners. *New Language Dimensions Journal of Literature, Linguistics, and Language Teaching, 5*(1).

Jewitt, C. (2009). *The Routledge handbook of multimodal analysis*. Routledge.

Jewitt, C. (2014). An introduction to multimodality. In C. Jewitt (Ed.), *The Routledge handbook of multimodal analysis* (2nd ed., pp. 15–30). Routledge.

Kaowiwattanakul, S. (2025). Multimodal literature and CEFR reading proficiency: Improving literary reading skills in EFL learners. *Indonesian Journal of Applied Linguistics, 15*(1), 20–34. https://doi.org/10.17509/ijal.v15i1.75921

Kress, G., & Jewitt, C. (2010). *Multimodality, literacy and school English*. Continuum.

Kress, G., & van Leeuwen, T. (2006). *Reading images: The grammar of visual design* (2nd ed.). Routledge. https://doi.org/10.4324/9780203619728

Lim, F. V., Toh, W., & Nguyen, T. T. H. (2022). Multimodality in the English language classroom: A systematic review of literature. *Linguistics and Education, 69*, 101048. https://doi.org/10.1016/j.linged.2022.101048

Mayer, R. E. (2001). *Multimedia learning*. Cambridge University Press.

Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). Cambridge University Press. https://doi.org/10.1017/CBO9781139164603

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). SAGE Publications.

Muhassin, M., Putri, N., Hidayati, D. A., & Pradana, S. A. (2022). A multimodal discourse analysis of English textbooks' covers. *EngEdu Journal, 15*(1). https://ejournal.radenintan.ac.id/index.php/ENGEDU

Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic analysis: Striving to meet the trustworthiness criteria. *International Journal of Qualitative Methods, 16*(1). https://doi.org/10.1177/1609406917733847

O'Cathain, A., & Thomas, K. J. (2004). "Any other comments?" Open questions on questionnaires: A bane or a bonus to research? *BMC Medical Research Methodology, 4*, Article 25. https://doi.org/10.1186/1471-2288-4-25

O'Halloran, K. L. (2004). *Multimodal discourse analysis: Systemic-functional perspectives*. Continuum.

Osborn, D. (2017). Constructing Israeli and Palestinian identity: A multimodal critical discourse analysis of world history textbooks and teacher discourse. *Journal of International Social Studies*, 4–33.

Quang Dong, L., Bich Ngoc, P., Binh Khiem, N., Quyen, N., & Phong, H. (2024). A multimodal discourse analysis of a medical textbook in Vietnam: Preliminary findings. *VNU Journal of Foreign Studies, 40*(3). https://doi.org/10.25073/2525-2445/vnufs.4908

Rowsell, J., & Walsh, M. (2011). Rethinking literacy education in new times: Multimodality, multiliteracies, and new literacies. *Brock Education Journal, 21*(1), 53–62.

Serafini, F. (2012). Reading multimodal texts in the 21st century. *Research in the Schools, 19*(1), 26–32.

Skinner, E. A., Kindermann, T. A., & Furrer, C. J. (2009). A motivational perspective on engagement and disaffection. *Educational and Psychological Measurement, 69*(3), 493–525. https://doi.org/10.1177/0013164408323233

Tseronis, A. (2018). Multimodal argumentation in educational materials. *Discourse & Society, 29*(3), 331–350. https://doi.org/10.1177/0957926517734667

Unsworth, L. (2024). Multimodal literacy in a new era of educational technology. *ECNU Review of Education, 7*(2), 384–405. https://doi.org/10.1177/20965311231179738

Unsworth, L., & Mills, K. A. (2020). English language teaching of attitude and emotion in digital multimodal composition. *Journal of Second Language Writing, 47*, 100712. https://doi.org/10.1016/j.jslw.2020.100712

Unsworth, L., & Ngo, T. (2021). Images and disciplinary literacy in secondary textbooks. *Visual Communication*. https://doi.org/10.1177/14703572211011169

Yeh, H.-C., & Lai, W.-C. (2019). Effects of multimodal learning on EFL students' reading comprehension and critical thinking. *Educational Technology & Society, 22*(1), 16–28.