



Multimodal Learning Module Based on Augmented Reality Technology to Enhance Students' Information and Digital Literacy

Modul Pembelajaran Multimodal Berbasis Teknologi Augmented Reality untuk Meningkatkan Literasi Informasi dan Digital Siswa

Stella Talitha, Yeti Mulyati, Yulianeta, Khaerudin Kurniawan

Indonesian Language Education Study Program Universitas Pendidikan Indonesia

stellatalitha@upi.edu

<https://doi.org/10.5614/sostek.itbj.2025.24.3.5>

Submitted: August 3, 2025 Accepted: November 24, 2025 Published: November 30, 2025

ARTICLE INFO

Keywords:

multimodality, modal integration, textbook analysis, content analysis

ABSTRACT

This study investigates the multimodal representation in the Grade 10 Bahasa Indonesia textbook published under the Kurikulum Merdeka, focusing on how various semiotic modes verbal, visual, typographic, layout, and color are functionally integrated to support meaning-making and student comprehension. Employing a qualitative content analysis method, data were collected through document analysis and analyzed using a multimodal categorization instrument. This instrument classified modes by type, communicative function, and modal integration. The analysis involved identifying modes on selected pages, evaluating the communicative function of each mode, and interpreting the relationships among modes, categorized as complementary, redundant, dominant, or incoherent, to assess cohesion and pedagogical relevance. The results show that visual modes are the most dominant, with typographic and layout features providing emphasis and structural clarity. Complementary relations were most frequent, demonstrating a high degree of synergy between modes in conveying meaning. Instances of redundancy, dominance, or incoherence were minimal but signal areas for improvement. The study concludes that the textbook reflects a pedagogical shift toward multimodal literacy aligned with 21st-century competencies. Practically, the findings inform educators and textbook designers on how to optimize visual-verbal coordination, design layout more coherently, and utilize typography and color meaningfully to foster more inclusive, engaging, and effective language instruction.

INFO ARTIKEL

Kata kunci:

multimodalitas, analisis buku teks, analisis isi, integrasi moda

ABSTRAK

Penelitian ini mengkaji representasi multimodal dalam buku teks Bahasa Indonesia kelas X yang diterbitkan dalam kerangka Kurikulum Merdeka, dengan fokus pada bagaimana berbagai moda semiotik (verbal, visual, tipografi, tata letak, dan warna) diintegrasikan secara fungsional untuk mendukung pemaknaan dan pemahaman siswa. Penelitian ini menggunakan metode analisis isi kualitatif, dengan pengumpulan data melalui studi dokumen dan dianalisis menggunakan instrumen kategorisasi multimodal. Instrumen ini mengklasifikasikan moda berdasarkan jenis, fungsi komunikatif, dan bentuk integrasi antarmoda. Proses analisis meliputi identifikasi moda pada halaman-

halaman terpilih, evaluasi fungsi komunikatif tiap moda, serta interpretasi hubungan antarmoda—dikategorikan sebagai hubungan komplementer, redundan, dominan, atau tidak koheren—untuk menilai kohesi dan relevansi pedagogis. Hasil penelitian menunjukkan moda visual paling dominan, sedangkan unsur tipografi dan tata letak berperan dalam penekanan dan kejelasan struktur informasi. Hubungan komplementer paling banyak ditemukan menunjukkan sinergi tinggi antarmoda dalam menyampaikan makna. Beberapa kasus hubungan redundan, dominan, dan tidak koheren ditemukan dalam jumlah kecil, namun mengindikasikan perlunya perbaikan dalam aspek tata letak dan koordinasi antarmoda. Penelitian ini menyimpulkan buku teks mencerminkan pergeseran pedagogis menuju literasi multimodal yang sejalan dengan kompetensi abad ke-21. Secara praktis, temuan ini memberikan panduan bagi pendidik dan perancang buku teks untuk mengoptimalkan koordinasi visual verbal, merancang tata letak yang lebih koheren serta memanfaatkan tipografi dan warna secara bermakna guna menciptakan pembelajaran bahasa yang lebih inklusif, menarik, dan efektif.

Introduction

Transformation in how individuals access, process, and construct meaning from information in the digital age has shifted educational paradigms, particularly in language learning. Communication is no longer limited to verbal modes (spoken or written) but increasingly involves various semiotic modes—such as visual, spatial, typographic, and color—that operate simultaneously to form multimodal meaning (Kress & van Leeuwen, 2006). In this context, multimodality refers to the use of two or more modes within a single text or communicative situation, resulting in richer and more complex representations of meaning. Kress (2010) emphasizes that each mode has its own representational potential and can complement one another in conveying messages. In today’s visually and digitally mediated world, learning must adapt to these contemporary forms of communication.

In literacy and language education, multimodality refers to the diverse ways of communicating and constructing meaning through various sign systems, such as text, images, sound, gesture, and spatial arrangement (Jewitt, 2008). Each mode contributes distinct representational strength and works synergistically in the learning context. In textbooks, multimodality appears through combinations of verbal narrative, illustrations, infographics, typography, color, and page layout—all designed to support deeper comprehension. Multimodal elements facilitate understanding, especially since not all students process written text optimally; some respond better to visual elements or rich visual-verbal combinations (Serafini, 2012).

Within a multimodal literacy approach, meaning is no longer interpreted linearly through written language alone but is constructed simultaneously across multiple channels of communication (Bezemer & Kress, 2010). Therefore, the selection and arrangement of modes in instructional materials become crucial design decisions, closely tied to reader identity, reading context, and intended learning goals. Textbooks that functionally integrate multimodality can increase student engagement, enrich learning experiences, and nurture critical literacy. In the context of Indonesian language instruction at the high school level, this means students are assessed not only for verbal comprehension but also for their ability to read visuals, interpret symbols, and link texts with context.

From a critical discourse analysis perspective, particularly within Fairclough’s framework, representations in educational texts are never neutral—they are ideologically charged and shaped by discursive choices. Fairclough (2003) argues that discourse serves as a means to represent the world, including social identities and relationships, and that such representations are embedded in broader power relations. Through lexical, grammatical, and visual strategies, textbooks may naturalize certain worldviews while marginalizing others. Fairclough (1992) also notes that language use in education serves not only to convey content but also to reproduce or challenge dominant ideologies through representational practices. These representations are formed through intertextual and interdiscursive

relations, which influence how learners engage with knowledge (Fairclough, 1995). Thus, analyzing multimodal representations in textbooks is not only a pedagogical concern but also a critical inquiry into how meaning and ideology are constructed.

Research by Rowsell and Walsh (2011) shows that students exposed to multimodal learning sources tend to demonstrate higher narrative capacity and creative expression. The use of multimodal elements also increases motivation and critical thinking, as learners are encouraged to interpret not only what a text says but also how it communicates. Multimodality in textbooks should therefore not be seen as mere decorative visuals but as integral to a student-centered instructional strategy. In this framework, analyzing multimodal representation in Indonesian language textbooks becomes increasingly relevant.

In the digital age, meaning-making in educational materials increasingly relies on the synergy of semiotic modes. Recent studies affirm that multimodality plays a crucial role in literacy development, particularly in language learning (Unsworth, 2020; Mills & Exley, 2021). Multimodal literacy goes beyond reading and writing; it includes the ability to interpret images, layout, typography, and digital media (Ting et al., 2020; Adami & Jewitt, 2021). In language textbooks, these elements often appear together, forming a complex textual environment that requires visual and critical literacy (Mei, 2022; Bezemer & Kress, 2020). Experts argue that well-integrated multimodal representation can enhance comprehension, support diverse learning styles, and promote critical thinking (Lim & Tan, 2021; Landa et al., 2023). However, the degree of multimodal integration and its pedagogical coherence still varies across curriculum contexts (Choi & Yi, 2019; Liu & Xu, 2020). In Indonesia, where curriculum reform moves toward student-centered approaches like Kurikulum Merdeka, multimodality in Indonesian language textbooks becomes an increasingly relevant research focus (Nuryani & Setiawan, 2022). This perspective highlights the importance of examining how visual, color, and layout modes are functionally coordinated with verbal content to support learning objectives.

Recent research further stresses the need to explore how textbooks meaningfully and pedagogically employ multimodal elements. Elmiana (2019) found that while visuals appear in Indonesian EFL textbooks, they are often decorative rather than instructional. Weninger (2020) argues that multimodality should be examined critically as part of textbook studies, since layout, images, and text collectively reflect deep ideological and pedagogical choices. In the Indonesian context, Prihatiningsih, Petrus, and Silvhiany (2021) show that cultural representation in junior high school EFL textbooks is visually limited and tends to reinforce mainstream norms while marginalizing local diversity. Similar findings by Ayu (2020) and Husain, Zuhri, and Musfirah (2021) indicate that multimodal content in textbooks often centers Western cultural imagery with minimal contextual relevance, risking a sense of cultural disconnection for students. A broader comparative view from Lee and Li (2020) reveals that multimodal representations in textbooks from China and Hong Kong not only shape national identity but also influence student perceptions of culture, underlining the need for critically designed multimodal learning resources.

As a primary medium of instruction, textbooks play a central role in developing students' literacy skills. In Indonesia, Bahasa Indonesia textbooks serve not only as teaching resources but also as tools for cultivating critical thinking and appreciation for language and culture. With the implementation of Kurikulum Merdeka, which emphasizes experiential and conceptual learning, the presence of multimodal representation becomes increasingly important to address diverse learning needs. However, the integration of multimodality in Indonesian language textbooks remains an area open for deeper and more critical investigation.

Previous studies show that although multimodal elements appear in textbook design, they are not always pedagogically integrated. Wijayanti (2020) noted that many illustrations in Bahasa Indonesia textbooks are decorative and do not reinforce meaning. Sari and Pratama (2021) observed mismatches between narrative text and illustrations in high school textbooks. Ningsih (2019) also highlighted the dominance of verbal and visual modes, with minimal attention to typography and layout as tools for understanding. On the other hand, more optimistic views from Cichocka (2016) and Madjid (2002) suggest that multimodality supports more contextual learning, especially in books developed under the

2013 Curriculum. These diverse findings suggest a gap in understanding how multimodal integration actually supports student comprehension.

To date, few studies have specifically examined multimodal representation in Bahasa Indonesia textbooks developed under the Kurikulum Merdeka, particularly regarding the communicative functions of each mode and their relevance to language learning. One such study by Nuryani and Setiawan (2022) analyzed the visual and textual integration in Kurikulum Merdeka textbooks but focused primarily on aesthetic features rather than pedagogical function. However, their research did not systematically evaluate how different modes interact to construct meaning in support of learning objectives. This gap highlights the need for further studies that not only describe the presence of multimodal elements but also assess their communicative roles and intermodal cohesion. By addressing this, the present study aims to offer a more comprehensive perspective on how multimodal integration contributes to the effectiveness of Bahasa Indonesia instructional materials.

This research aims to analyze multimodal representation in Indonesian language textbooks at the high school level. Specifically, it identifies the types of modes used—verbal, visual, typographic, layout, and color—and evaluates how these modes interact to construct meaning and support student understanding. The novelty of this study lies in its integrative approach, combining multimodal analysis with a critical evaluation of content delivery in contemporary Bahasa Indonesia textbooks. The findings are expected to contribute to the development of more communicative, functional, and learner-centered instructional materials aligned with 21st-century educational needs.

Method

This study employs a qualitative approach using content analysis to examine multimodal representation in Bahasa Indonesia textbooks used at the senior high school (SMA) level. This approach was chosen because it allows the researcher to deeply understand and interpret meanings constructed through the combination of various semiotic modes (Moleong, 2018). Content analysis is considered appropriate for identifying and evaluating multimodal elements in texts, including verbal content, images, typography, color, and page layout (Hermawan, 2013).

The subjects of this study are two Bahasa Indonesia textbooks for Phases E and F at the senior high school level, developed under the Kurikulum Merdeka and published by the Ministry of Education, Culture, Research, and Technology of the Republic of Indonesia. These textbooks were selected due to their widespread use as reference materials in the implementation of the latest curriculum across various educational institutions (Hartati, Sukenti, & Nazirun, 2024).

The research procedure began with selecting units of analysis in the form of pages or chapters that prominently feature combinations of verbal and visual texts. These materials were analyzed as complete textual units containing semiotic modes such as textual narratives, illustrations, page layout, typography, and color use. A multimodal categorization sheet, developed based on the theoretical framework of Kress and van Leeuwen (2006), was used as an instrument to classify types of modes, their communicative functions, and the degree of their integration in meaning-making.

Data were collected through document analysis by thoroughly reading and annotating all multimodal elements found on the selected pages. The data were then analyzed using a descriptive-interpretative approach through four stages: (1) identifying the types of modes used; (2) classifying the communicative functions of each mode within the learning context; (3) analyzing intermodal cohesion to evaluate each mode's contribution to meaning-making; and (4) interpreting the potential impact of multimodal integration on student comprehension.

To ensure data validity, the study employed theoretical triangulation as proposed by Patton (1999), by comparing findings across different theoretical perspectives, particularly multimodal literacy, semiotic theory, and critical discourse analysis. The conclusions were drawn through interpretive analysis, supported by recurring patterns in the data and cross-referenced with relevant literature. In addition,

expert validation was conducted through consultations with three specialists in the fields of literacy and language education—two senior lecturers in applied linguistics and one curriculum expert. Their input was used to evaluate the relevance, clarity, and accuracy of the multimodal function classifications and to assess the appropriateness of the intermodal analysis framework in relation to pedagogical practices.

Results and Discussion

This study aims to examine multimodal representations in the Grade 10 Bahasa Indonesia textbook using a multimodal analysis instrument. The analysis of multimodal representations in textbooks is critical in the context of 21st-century literacy, which requires learners not only to comprehend verbal texts but also to interpret visual elements, layout, typography, and other graphic features. Therefore, the presence of diverse semiotic modes in the textbook must be assessed in terms of their occurrence, communicative function, and cohesion in supporting content comprehension.

The multimodal analysis in this study was conducted using an instrument (Table I) that outlines five principal modes: verbal, visual, typographic, layout, and color. Each mode was identified based on specific indicators and evaluated according to its role in constructing and enhancing textual meaning. This approach enabled the researchers to determine the extent to which multimodality is functionally integrated within the textbook and how each mode contributes to achieving communicative and pedagogical goals across various instructional units.

Multimodal Analysis of the Textbook

The identification results from the Bahasa Indonesia Grade 10 textbook reveal that multimodal representations in this textbook encompass five primary modes: verbal, visual, typographic, layout, and color. In total, sixteen multimodal entries were found to meet the identification indicators and communicative functions as outlined in the analysis instrument. These findings indicate that the multimodal elements in the textbook serve not merely aesthetic purposes but are intentionally designed to enrich meaning and support the learning process.

The distribution of the identified modes can be detailed as follows:

- Verbal Mode: 2 entries (negotiation narrative and explanatory text on plastic waste),
- Visual Mode: 8 entries (illustrations, infographics, comics, historical photographs, and concept maps),
- Typographic Mode: 2 entries (use of capital letters and font variations for emphasis),
- Layout Mode: 2 entries (arrangement of text and images in observation reports and comic strips),
- Color Mode: 2 entries (use of color in infographics and page backgrounds).

These findings are presented systematically in Table I below to provide a detailed overview of the types of modes, their communicative functions, and examples of multimodal representations identified in the Bahasa Indonesia Grade 10 textbook.

Table I Identified Types of Multimodal Representations in the Bahasa Indonesia Grade 10 Textbook

No	Type of Mode	Indicator	Communicative Function	Example	Page
1	Verbal	Narrative text	Conveying the main topic	Negotiation text between Irfan and the seller	p. 203
2	Verbal	Expository text	Delivering factual information	Explanation on plastic waste	p. 74
3	Visual	Illustration of grasshopper	Describing an observed object	Illustration of orchid mantis	p. 28

No	Type of Mode	Indicator	Communicative Function	Example	Page
4	Visual	Infographic of national parks	Supporting environmental description	Infographic on Indonesian national parks	p. 31
5	Visual	Image of Bosscha Observatory	Describing a research site	Bosscha Observatory photograph	p. 29
6	Visual	Comic strip “Colored Pencils”	Illustrating social critique	Comic on bullying	p. 66
7	Visual	Infographic on cyberbullying	Reinforcing social data	Infographic on bullying impact	p. 72
8	Visual	Infographic on smoking risks	Explaining health impact	Smoking hazard chart	p. 76
9	Visual	Photo of Hikayat manuscript	Enhancing historical content	Illustration of <i>Hikayat Bayan Budiman</i>	p. 90
10	Visual	Concept map of short story	Mapping narrative structure	Visual structure of a short story	p. 107
11	Typography	Chapter title in capital letters	Drawing attention	Chapter I–VI titles	Throughout the textbook
12	Typography	Font variation in infographic	Emphasizing key information	Titles and data in smoking infographic	p. 76
13	Layout	Text-image arrangement in report	Guiding reading flow	Observation report layout	pp. 28–31
14	Layout	Comic panel placement	Structuring visual narrative	Comic “Colored Pencils”	p. 66
15	Color	Infographic background color	Differentiating data categories	National park infographic	p. 31
16	Color	Gradient color in health chart	Providing visual emphasis	Smoking hazard infographic	p. 76

Table I reveals that the most dominant type of mode found in the Bahasa Indonesia Grade 10 textbook is the visual mode, with a total of eight identified instances. Visual elements—such as illustrations, photographs, infographics, comics, and concept maps—are employed to clarify, enrich, and support the content of verbal texts presented across various instructional units. The predominance of visual modes suggests that the textbook does not rely solely on verbal explanations but actively utilizes visual support to enhance students’ comprehension of the discussed concepts and phenomena. Alongside visual elements, verbal, typographic, layout, and color modes also play key roles in reinforcing the informational structure and enriching students’ multimodal learning experience.

Verbal modes were identified twice, primarily in the form of narrative and expository texts that form the backbone of the book’s instructional content. Similarly, typographic, layout, and color modes each appeared twice, indicating a deliberate effort to employ non-verbal visual elements such as font size, page organization, and color selection to guide reader focus, establish emotional tone, and improve navigability. These findings underscore the importance of adopting a multimodal approach in modern textbook design, in line with the principles of multimodal literacy that integrate multiple semiotic resources to enhance the learning process (Kress & van Leeuwen, 2006; Serafini, 2012).

To further illustrate the proportion of modality types found in the Bahasa Indonesia Grade 10 textbook, Figure 1 presents a pie chart visualizing the distribution of the five identified modes. This diagram clearly demonstrates the dominance of visual representation, followed by verbal, typographic, layout, and color modes in smaller proportions. Such visualization enables readers to quickly grasp the overall tendencies of multimodal integration applied in the textbook’s composition.

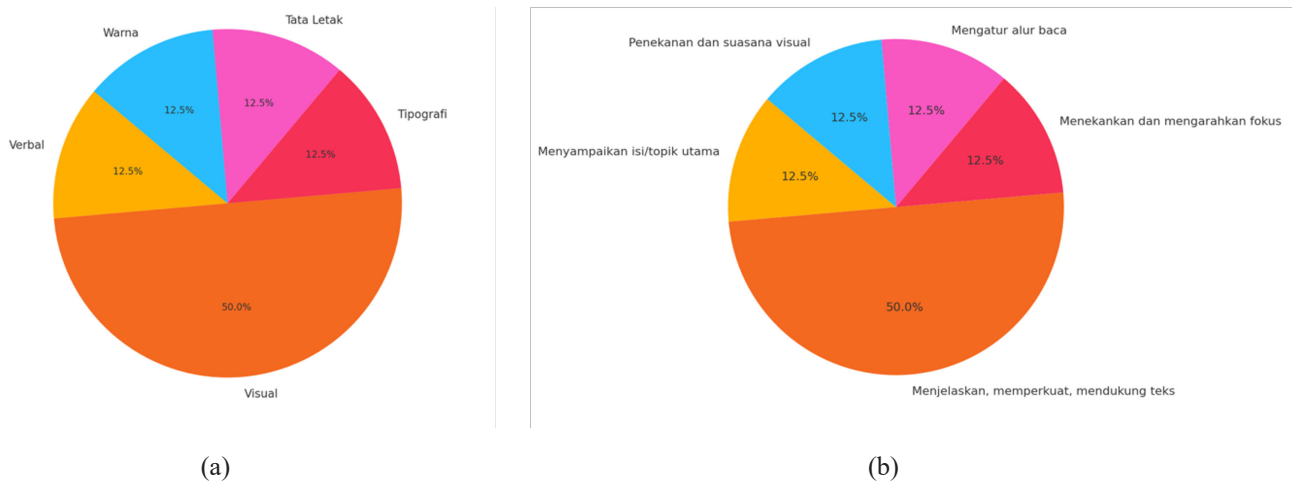


Figure 1 Distribution of Types of Modalities in the Bahasa Indonesia Grade 10 textbook. (a) Visual representation of the proportion of verbal, visual, typographic, layout, and color modalities; (b) Visual modality dominates among the identified multimodal representations.

1. Verbal Mode

The textbook employs verbal mode primarily through narrative and expository texts, such as the negotiation dialogue (p. 203) and plastic waste explanation (p. 74). Though only two entries are listed in the multimodal table, the book is inherently text-based. Verbal texts provide the linguistic backbone for multimodal synergy, facilitating both communicative and cognitive engagement within realistic contexts.

2. Visual Mode

Visuals are the most dominant mode, with eight entries including illustrations (orchid mantis, p. 28), photographs (Bosscha Observatory, p. 29), infographics (national parks, p. 31), comics (Colored Pencils, p. 66), and concept maps (p. 107). These visuals clarify verbal information and stimulate both cognitive and emotional responses, aligning with Paivio's dual coding theory and reflecting a broader shift toward image-based literacy in digital-native learning environments.

3. Typographic Mode

Typography appears in chapter titles (capital letters) and within infographics (font variation on p. 76). It structures content hierarchically and draws attention to key messages. According to Kress and van Leeuwen (2006), typography shapes reader perception and acts as a multimodal resource that enhances emphasis and readability.

4. Layout Mode

Layout structures content by spatially arranging verbal and visual elements. In the observational report (pp. 28–31), texts and images are parallel; in the comic (p. 66), panels follow a clear narrative path. Serafini (2012) notes that layout supports intuitive reading and reduces cognitive load, aiding student comprehension through spatial logic.

5. Color Mode

Color carries symbolic and emphatic functions beyond decoration. It helps differentiate data categories (p. 31) and emphasize content severity (p. 76). Burmark (2004) highlights its role in visual memory and reader engagement. Color use enhances emotional tone and clarifies content meaning within multimodal learning materials.

Analysis of Modal Integration

The analysis of modal integration was conducted using the instrument presented in Table 2, which classifies the relationships among modes into five categories: complementary, redundant, verbal-dominant, visual-dominant, and incoherent. These categories enabled the study to evaluate whether the modes presented in the textbook work synergistically, deliver overlapping information, or whether one mode dominates the meaning-making process—potentially creating dissonance. This classification provides a framework for assessing the extent to which multimodal integration is functionally and effectively implemented in the textbook.

This section presents the results of the modal integration analysis based on the previously identified data, followed by a scholarly interpretation of the relationships among modes. The findings are expected to contribute to evaluating the quality of multimodality in the Bahasa Indonesia textbook and to serve as a foundation for developing more pedagogically sound multimodal learning materials in the future.

Table II outlines the results of modal integration analysis in the Grade 10 Bahasa Indonesia textbook using the relational classification instrument. Modal integration is a critical aspect to ensure that verbal, visual, typographic, layout, and color elements not only coexist on the page but also collaborate meaningfully to construct coherent messages that support the learning process. In this study, each multimodal combination was examined to determine the type of relationship it represented, as categorized under the five relational types: complementary, redundant, verbal-dominant, visual-dominant, and incoherent.

Table II Modal Integration Analysis in The *Bahasa Indonesia* Grade 10 Textbook

No	Modal Relation	Description of Relation	Example in the Textbook	Page
1	Complementary (C1)	Visual mode clarifies verbal content	Illustration of a grasshopper clarifies the observation report text	p. 28
2	Complementary (C1)	Illustration of Bosscha Observatory supports the description of the site	Visualization of the astronomical site in the observation report	p. 29
3	Complementary (C1)	Comic on bullying reinforces editorial text	“Colored Pencils” comic illustrates a narrative of social critique	p. 66
4	Redundancy (C2)	Infographic on plastic waste reiterates verbal explanation	Plastic waste infographic and expository text present the same information	p. 74
5	Redundancy (C2)	Smoking infographic conveys the same data as the text	Smoking hazard chart complements the argumentative explanation	p. 76
6	Verbal-Dominant (C3)	Verbal explanation of plastic waste dominates; image is merely illustrative	The text is more informative; image serves only decorative function	p. 74
7	Visual-Dominant (C4)	Concept map of short story emphasizes structure more than verbal explanation	Visual diagram of short story structure is more prominent than verbal description	p. 107
8	Complementary (C1)	Manuscript photo enriches historical text	Image of the classical manuscript enhances the <i>Hikayat</i> narrative	p. 90
9	Complementary (C1)	Collage of poets strengthens poetry learning	Images of Indonesian poets support the poetry material	p. 223
10	Incoherent (C5)	Layout of infographic on national parks is inconsistent with accompanying text	Image placement does not follow the logical sequence of the text	p. 31

Table II indicates that complementary relations (C1) are the most dominant type of modal integration in the Bahasa Indonesia Grade 10 textbook. Six data points demonstrate that verbal and visual modes work harmoniously to construct meaning. In such relations, illustrations, photographs, or diagrams serve to clarify and enrich the information conveyed through verbal text, reinforcing the connection between textual content and students’ visual experiences. The predominance of complementary modes suggests that the textbook applies fundamental principles of effective multimodal literacy, where meaning is

constructed through the synergy of various semiotic resources (Kress & van Leeuwen, 2006; Serafini, 2012).

In addition to complementary relations, two instances of redundant relations (C2) were identified, in which verbal and visual elements present the same information to reinforce a message—for example, in the infographics on plastic waste and the dangers of smoking. Dominance relations were found in only one instance each: verbal-dominant (C3) and visual-dominant (C4). This indicates that in a small number of cases, one mode carried more informational weight than the others. A single case of incoherence (C5) was also noted, suggesting that not all multimodal integrations were implemented flawlessly and that improvements in layout and mode coordination remain necessary in some sections. Overall, this integration pattern reflects an intentional effort to enhance learning effectiveness through multimodal textbook design, though minor refinements are still needed to achieve optimal multimodal cohesion.

To clarify the distribution of modal relations identified in the Bahasa Indonesia textbook, a pie chart is presented. This visual illustrates the proportions of each relation type—complementary, redundant, verbal-dominant, visual-dominant, and incoherent—based on the frequency of occurrence in the analyzed data. The graphic presentation aims to provide a quick overview of the multimodal integration quality in the textbook, highlighting both its strengths and areas for improvement in multimodal design implementation.

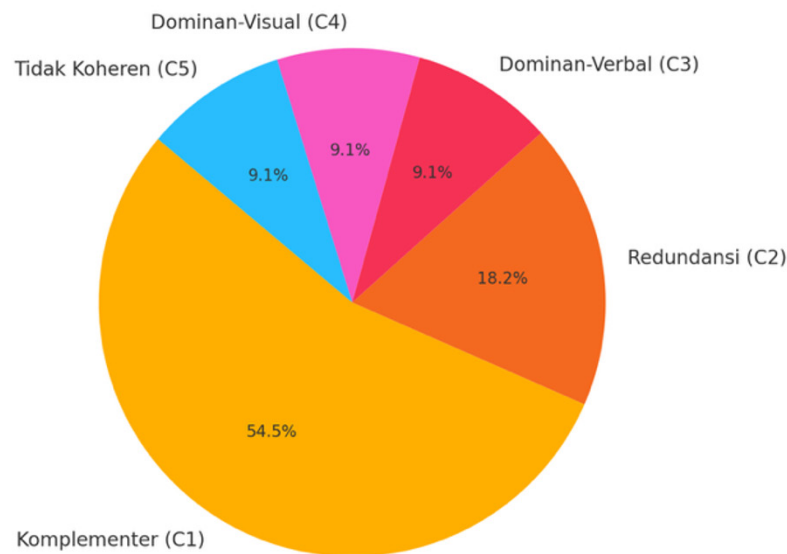


Figure 2 Distribution of Modal Cohesion Types in the Bahasa Indonesia Grade 10 textbook. (a) Visualization of the proportion of Complementary, Redundancy, Dominant-Verbal, Dominant-Visual, and Incoherent relations; (b) Complementary relations are the most dominant among the identified multimodal interactions.

The integration of multiple semiotic modes within a single textbook reflects not only aesthetic choices but deliberate pedagogical decisions designed to support comprehension, engagement, and learning effectiveness. In the case of the Grade 10 Bahasa Indonesia textbook, the analysis of modal integration reveals the extent to which verbal, visual, typographic, layout, and color elements function in unison—or tension—to construct coherent, engaging, and pedagogically meaningful instructional content.

1. Complementary Modal Integration (C1)

The most frequent integration, with six examples including the grasshopper illustration (p. 28), Bosscha Observatory photo (p. 29), and “Colored Pencils” comic (p. 66). These visuals enhance verbal meaning, enabling multimodal scaffolding (Cope & Mayer, 2009). Each mode supports meaning-making through distinct yet synergistic functions.

2. **Redundancy (C2)**
Seen in two infographics: plastic waste (p. 74) and smoking hazards (p. 76). Verbal and visual modes present the same information, reinforcing key points. Though pedagogically simpler than C1, redundancy aids retention and supports inclusive learning styles (Sweller, 2010).
3. **Verbal-Dominant (C3)**
One case on plastic waste (p. 74) where text carries most of the meaning and the visual is minimal. This reflects a conventional model, underutilizing multimodal potential for engagement or emotional impact.
4. **Visual-Dominant (C4)**
Identified once with the short story concept map (p. 107). The visual explains narrative structure more clearly than the text. This promotes visual literacy and metacognitive skill development, especially for abstract content.
5. **Incoherent (C5)**
One incoherent case in the national parks infographic (p. 31), where layout disrupts textual alignment. As Bezemer & Kress (2010) highlight, poor modal coordination undermines comprehension and should be addressed in future revisions.

Visualizing Modal Integration: Trends and Implications

To offer a concise overview of the modal integration landscape in the textbook, Figure 2 illustrates the distribution of the five modal relationship types. Complementary relations dominate (60%), followed by redundancy (20%), and single instances of dominance (verbal and visual) and incoherence.

This distribution affirms that the textbook's multimodal design is largely intentional and pedagogically sound. Complementary integration—where different modes reinforce and enrich one another—is the ideal in multimodal instructional design. Its prevalence in this textbook suggests a growing awareness among curriculum designers and content developers about the value of multimodality in language education.

However, the appearance of redundancy, dominant modes, and incoherence also indicates areas for further refinement. Ideally, multimodal resources should neither repeat content unnecessarily nor privilege one mode to the detriment of others—unless pedagogically justified. A balanced integration across modes offers learners diverse entry points to meaning, supports differentiated instruction, and caters to various cognitive styles.

Pedagogical Significance of Modal Integration

The implications of these findings go beyond textbook evaluation. They offer insights into how multimodal integration shapes students' learning experiences in Bahasa Indonesia classrooms. When modes are harmonized:

1. Students engage both analytical and intuitive faculties (dual coding).
2. Complex concepts are scaffolded across representations.
3. Affective and social dimensions of learning are activated.
4. Learning materials become more inclusive and equitable.

In contrast, modal imbalance or incoherence can create barriers to understanding, particularly for students with different learning needs. Thus, analyzing modal integration is not merely a theoretical exercise but a crucial part of inclusive, effective curriculum development.

Results and Discussion

This study was conducted to address the central question posed in the introduction: to what extent are multimodal representations functionally integrated in the Bahasa Indonesia Grade 10 textbook, and how does modal cohesion contribute to meaning-making within the context of language learning at the senior high school level? The analysis reveals that the textbook incorporates five main types of modes—verbal, visual, typographic, layout, and color—with the visual mode appearing most frequently. This supports the initial hypothesis that contemporary textbooks, particularly those developed under the Kurikulum Merdeka framework, have begun to adopt a more explicit and deliberate multimodal approach. Similar trends have been found by Unsworth (2020), who noted an increase in multimodal content in English language textbooks in Australia, and by Adami and Jewitt (2021), who observed a global shift toward multimodal pedagogy in literacy education.

Scientifically, these findings reflect a paradigm shift in instructional material development—from a predominantly verbal-centric approach to a more holistic multimodal perspective. Visual modes such as illustrations, comics, photographs, and infographics are not merely decorative; they reinforce textual meaning, visualize complex data, and clarify abstract concepts. Serafini (2012) emphasized that visual elements serve as interpretive resources that actively contribute to meaning-making. This interpretation aligns with the “reading path” theory of Kress and van Leeuwen (2006), which emphasizes the importance of cohesive modal integration in constructing a coherent and meaningful reading experience.

The modal integration analysis presented in Table II further confirms that complementary relations (C1) are the most prevalent. In the majority of cases, verbal and visual modes work together to deliver meaning. This echoes findings from Lim and Tan (2021), who demonstrated that textbooks with strong verbal-visual synergy improve student comprehension and engagement. The presence of redundant relations (C2), though less frequent, also suggests a pedagogical intention to reinforce key messages through repetition—an approach supported by Cope & Mayer’s (2009) theory of dual coding, which posits that repetition across modalities can aid memory retention and processing.

In contrast, verbal-dominant (C3) and visual-dominant (C4) relations were each observed only once, indicating a relatively balanced integration of modes without overreliance on a single channel. This balance aligns with findings from Choi and Yi (2019), who argue that multimodal imbalance—particularly visual overload—can reduce instructional clarity. A single instance of incoherent relation (C5), involving misaligned infographic placement, highlights that while the integration is generally strong, certain areas of layout and mode coordination could benefit from refinement. Similar concerns were raised by Elmiana (2019), who found that some Indonesian EFL textbooks featured visuals with inconsistent alignment, affecting coherence.

When compared to earlier studies, these findings suggest progress in multimodal textbook design. Research by Wijayanti (2020) and Sari and Pratama (2021) criticized the excessive use of decorative illustrations in Bahasa Indonesia textbooks. In contrast, this study found that visual modes in the current textbook carry substantial representational functions and are meaningfully aligned with verbal texts. Furthermore, modes like typography and color—often overlooked in previous research—are identified here as purposeful tools for emphasis and reader navigation. This is in line with Lee and Li’s (2020) comparative study, which emphasized the role of layout and typographic choices in building national identity and cultural context in Asian textbooks.

The complementary mode, in particular, plays a vital role in helping students connect with the content emotionally and cognitively. For example, the image of the Bosscha Observatory supports the verbal explanation of the scientific text, enabling students to better visualize and comprehend the subject. The concept map of short story structure, though categorized as visual-dominant, effectively scaffolds students’ understanding of literary elements, which may be abstract or difficult to articulate verbally. Rowsell and Walsh (2011) found similar benefits in classrooms that employed multimodal texts, noting improvements in both narrative comprehension and creative engagement.

Redundant modes also serve a functional purpose by enhancing memorability and reinforcing concepts. The infographic on the dangers of smoking, for example, repeats data available in the verbal text but presents it visually to increase its impact and clarity. While redundancy may appear pedagogically less sophisticated than complementary integration, it supports differentiated learning styles and promotes retention. This echoes findings by Ting et al. (2020), who observed that redundancy in multimodal designs can support inclusive pedagogy and accommodate diverse learner preferences.

Moreover, typographic and color modes, while each identified only twice in the dataset, contribute to textual emphasis and affective engagement. Capital letters in chapter titles help organize information hierarchically, while font variation in infographics signals important data points. Color use, such as gradient backgrounds and contrasting tones in infographics, draws attention to critical issues and helps differentiate data categories. Studies by Burmark (2004) and Ningsih (2019) support this claim, emphasizing how typography and color improve information hierarchy and emotional resonance in learning materials.

The combination of these modalities—and the functional integration thereof—demonstrates a meaningful application of multimodal literacy principles in textbook development. It reflects an awareness of how students process information and the need for diverse entry points into text comprehension. Landa et al. (2023) argue that multimodal integration fosters reflective learning and empowers students to construct meaning actively.

Pedagogically, the implications of this study are significant. Teachers can use the results to better harness the textbook's multimodal features in classroom instruction. For instance, rather than focusing solely on verbal texts, educators might prompt students to interpret infographics, discuss the narrative in comics, or analyze how layout and design contribute to meaning-making. This integrated approach aligns with critical literacy practices that position students not just as readers but as interpreters and critics of multimodal texts (Unsworth & Mills, 2021; Weninger, 2020).

In curriculum design, the results affirm that multimodality should be embedded at the foundational level. Rather than treating visual and design elements as ancillary to content, curriculum developers and textbook writers must recognize them as integral components of meaning construction. This is particularly relevant in language education, where understanding nuance, tone, and rhetorical structure often depends on more than just verbal clues. Adami and Kress (2010) emphasize the necessity of multimodal awareness in developing 21st-century communicative competence.

Methodologically, this study benefits from its dual-layer analysis: Table I examines mode types and functions, while Table II investigates their interaction and cohesion. This approach offers both breadth and depth in evaluating multimodal integration. It allows researchers to not only catalog modal instances but also assess their relational quality—a necessary step for moving beyond descriptive toward critical textbook analysis (Serafini, 2015; Bezemer & Kress, 2020).

Another methodological strength lies in the study's alignment with the Kurikulum Merdeka framework, which emphasizes student-centered learning, critical thinking, and flexibility. The findings suggest that the analyzed textbook embodies these principles through its multimodal structure. However, the study also identifies room for improvement, particularly in layout coherence and the avoidance of decorative redundancy that lacks pedagogical value—concerns also noted by Prihatiningsih, Petrus, and Silvhiany (2021) in their critique of Indonesian EFL textbooks.

Conclusion

This study concludes that the Grade 10 Bahasa Indonesia textbook developed under the Kurikulum Merdeka demonstrates purposeful and functional multimodal integration, particularly through five primary modes: verbal, visual, typographic, layout, and color. The prominence of visual modes signals a paradigm shift from traditional text-centric instruction toward a more contextual and image-rich approach that facilitates comprehension through multiple pathways of meaning-making. The analysis of

modal integration shows that complementary relations are the most prevalent, indicating strong synergy between verbal and visual elements in shaping understanding. Instances of redundancy, dominance, and incoherence were also identified—though limited in number—reflecting a generally balanced and design-conscious application of multimodal principles.

The state-of-the-art contribution of this study lies in its dual-layered analytical approach, which not only identifies the presence of multimodal elements but also evaluates the quality of their intermodal interactions. By focusing on textbooks designed within the relatively underexplored Kurikulum Merdeka context, this research offers a comprehensive view of how multimodal literacy is being implemented in Indonesian high school language education. Pedagogically, the findings affirm the importance of modal synergy as a foundation for fostering critical and inclusive literacy—accommodating diverse learning styles and enriching students' cognitive and affective learning experiences.

From a methodological perspective, the analytical instrument developed in this study may be adapted for evaluating textbooks across disciplines and educational levels. It also provides a practical reference for curriculum developers and instructional material designers aiming to incorporate multimodality meaningfully. Further research should explore how students engage with multimodal elements in real classroom settings, including how they interpret visuals, navigate layout, or respond emotionally to color and design. Longitudinal studies could also be conducted to examine how multimodal resources affect students' literacy development over time. For textbook authors and curriculum planners, this study underscores the need to ensure meaningful intermodal coordination and to move beyond decorative visuals. In this way, multimodality becomes not merely a design trend but a foundational strategy for 21st-century learning—one that deepens understanding and invites active student engagement.

References

- Adami, E., & Jewitt, C. (2021). Multimodal literacy: From theories to practices. *Visual Communication, 20*(4), 453–468. <https://doi.org/10.1177/14703572211021369>
- Ayu, M. (2020). Evaluation of cultural content on English textbook used by EFL students in Indonesia. *Journal of English Teaching, 6*(3), 183–192. <https://doi.org/10.33541/jet.v6i3.1925>
- Bezemer, J., & Kress, G. (2010). *Multimodality, learning and communication: A social semiotic frame*. Routledge.
- Bezemer, J., & Kress, G. (2020). Social semiotics and multimodal texts in education. *Discourse, Context & Media, 35*, 100391. <https://doi.org/10.1016/j.dcm.2020.100391>
- Choi, Y., & Yi, Y. (2019). Teacher support of multimodal composition in multilingual classrooms. *Journal of Adolescent & Adult Literacy, 62*(6), 663–672. <https://doi.org/10.1002/jaal.935>
- Cichocka, A. (2016). Multimodal analysis of educational materials: Images in a history textbook. *Journal of Language and Education, 2*(2), 6–16.
- Cope, B., & Mayer, K. (2009). Multiliteracies: New literacies, new learning. *Pedagogies: An International Journal, 4*(3), 164–195. <https://doi.org/10.1080/15544800903076044>
- Elmiana, D. S. (2019). Pedagogical representation of visual images in EFL textbooks: a multimodal perspective. *Pedagogy, Culture & Society, 27*(4), 613–628. <https://doi.org/10.1080/14681366.2019.1569550>
- Hartati, S., Sukenti, D., & Nazirun. (2024). Penguatan literasi dalam Kurikulum Merdeka: Telaah buku teks Bahasa Indonesia. *Jurnal Pendidikan Bahasa dan Sastra Indonesia, 13*(1), 45–58.
- Hermawan, B. (2013). Multimodality in English language teaching textbooks: A case study of a textbook for junior high school students in Indonesia. *Indonesian Journal of Applied Linguistics, 2*(2), 65–72. <https://doi.org/10.17509/ijal.v2i2.70>
- Jewitt, C. (2008). Multimodality and literacy in school classrooms. *Review of Research in Education, 32*(1), 241–267. <https://doi.org/10.3102/0091732X07310586>

- Kress, G. (2010). *Multimodality: A social semiotic approach to contemporary communication*. Routledge.
- Kress, G., & van Leeuwen, T. (2006). *Reading images: The grammar of visual design* (2nd ed.). Routledge.
- Husain, N. F., Zuhri, M. D., & Musfirah, M. (2021). Semiotic analysis to the cultural content in the EFL textbook. *International Journal of Research on English Teaching and Applied Linguistics*, 1(2), 67-82. <https://doi.org/10.30863/ijretal.v1i2.1230>
- Landa, R., Galvis, H., & Gutierrez, C. (2023). Integrating multimodal learning strategies in textbook design: A Latin American perspective. *International Journal of Educational Research Open*, 4, 100234. <https://doi.org/10.1016/j.ijedro.2023.100234>
- Lee, J. F. K., & Li, X. (2020). Cultural representation in English language textbooks: A comparison of textbooks used in mainland China and Hong Kong. *Pedagogy, Culture & Society*, 28(4), 605-623. <https://doi.org/10.1080/14681366.2019.1681495>
- Lim, F. V., & Tan, K. H. (2021). Multimodal literacy: The case of the Singapore English curriculum. *Curriculum Perspectives*, 41(2), 123–135. <https://doi.org/10.1007/s41297-021-00145-w>
- Liu, Y., & Xu, B. (2020). Multimodal representation in EFL textbooks: A systemic functional perspective. *System*, 89, 102200. <https://doi.org/10.1016/j.system.2019.102200>
- Madjid, A. (2002). Kontekstualisasi kurikulum bahasa Indonesia berbasis kompetensi. *Jurnal Bahasa dan Seni*, 30(2), 133–144.
- Mei, Y. (2022). The visual grammar of textbooks: Multimodal strategies in Chinese high school English books. *Linguistics and Education*, 69, 101009. <https://doi.org/10.1016/j.linged.2022.101009>
- Mills, K. A., & Exley, B. (2021). Multiliteracies and multimodal texts in classrooms. *The Reading Teacher*, 75(2), 209–217. <https://doi.org/10.1002/trtr.2003>
- Moleong, L. J. (2018). *Metodologi penelitian kualitatif* (Edisi revisi). PT Remaja Rosdakarya.
- Ningsih, D. A. (2019). Keterpaduan visual dan teks dalam buku pelajaran Bahasa Indonesia SMA. *Jurnal Kajian Linguistik dan Sastra*, 4(1), 73–84.
- Nuryani, D., & Setiawan, A. (2022). Exploring multimodality in Bahasa Indonesia textbooks: A case of Kurikulum Merdeka. *Journal of Language and Literacy Education*, 18(3), 55–68.
- Prihatiningsih, F., Petrus, I., & Silvhianny, S. . (2021). Cultural Representation in EFL Textbooks for the Seventh Graders: A Multimodal Analysis. *Lingua Cultura*, 15(1), 121–133. <https://doi.org/10.21512/lc.v15i1.7319>
- Rowell, J., & Walsh, M. (2011). Rethinking literacy education in new times: Multimodality, multiliteracies, and new literacies. *Brock Education Journal*, 21(1), 53–62.
- Sari, R. P., & Pratama, F. D. (2021). Peran ilustrasi dalam pembelajaran literasi: Studi terhadap buku teks SMA. *Jurnal Pendidikan dan Kebudayaan*, 26(3), 352–365.
- Serafini, F. (2012). Reading multimodal texts in the 21st century. *Research in the Schools*, 19(1), 26–32.
- Ting, Y. L., Tan, C. L., & Liew, W. L. (2020). Visual literacy and textbook evaluation: A multimodal perspective. *Asian Journal of English Language Studies*, 8(1), 87–104.
- Weninger, C. (2020). Multimodality in critical language textbook analysis. *Language, Culture and Curriculum*, 34(2), 133–146. <https://doi.org/10.1080/07908318.2020.1797083>
- Wijayanti, R. (2020). Analisis visual dalam buku teks Bahasa Indonesia kelas X. *Jurnal Ilmu Pendidikan*, 26(2), 180–191.