

A Multimodal Analysis of Meaning-Making in the “Train of Thought” English Textbook for Grade XII: A Descriptive Qualitative Study

Meysha Sandra Kirani Putri^{1*}, Dwi Andriani², Ainur Rohmah³, and Hastuti Retno kuspiyah⁴

^{1,2,3,4} UUniversitas Nurul Huda, Indonesia

* E-mail: meyshasandram@gmail.com

Abstract

This study aims to analyze multimodal meaning-making in the Train of Thought English textbook for Grade XII students by focusing on three key aspects: expression, interaction, and composition. Employing a descriptive qualitative research design, this study applies Kress and van Leeuwen's (2021) visual grammar framework to examine how visual and verbal elements contribute to the construction of meaning. The data were collected from selected pages of the textbook and analyzed using qualitative content analysis to explore the relationships among various semiotic resources. The findings indicate that multimodality is realized through the integration of text, images, color, and layout, which work together to support meaning-making. In terms of expression (representational metafunction), the textbook predominantly presents participants, actions, and concepts that facilitate students' understanding of the content. The interaction (interactional metafunction) is reflected through gaze, gestures, and linguistic features that engage readers and establish interpersonal relationships. Meanwhile, composition (compositional metafunction) organizes information through spatial arrangement, salience, and balance, guiding readers in interpreting the content effectively. Overall, the three elements—expression, interaction, and composition—are interconnected and play a significant role in enhancing students' comprehension and engagement. This study concludes that the Train of Thought textbook effectively implements multimodal principles to support English language learning, and the findings are expected to provide valuable insights for teachers, material developers, and researchers in designing and evaluating multimodal instructional materials.

Key words: Multimodality, Meaning-Making, Expression, Interaction, Composition.

INTRODUCTION

English has become a global language that connects people across borders and serves as a primary medium for communication, education, science, and technology. In Indonesia, English is a compulsory subject taught from secondary to tertiary education. As one of the main instructional resources, textbooks play a crucial role in facilitating learning by organizing materials, guiding classroom activities, and supporting the development of students' linguistic and communicative competence. A well-designed textbook not only provides language input but also enables students to construct meaning through various forms of representation. However, traditional textbooks have largely focused on written text and grammatical explanations, whereas current educational demands require more interactive, visually rich, and meaningful learning materials that reflect diverse communicative practices. English learning materials in Indonesian classrooms should integrate visual and textual elements to enhance students' comprehension and engagement (Sari, 2021). Similarly, Pratama and Wahyuni (2022) found that textbooks that incorporate multimodal elements, such as images and layout, can significantly improve students' participation and understanding, particularly at the secondary school level.

In this context, multimodality has become an essential concept in language education. Kress and van Leeuwen (2021) define multimodality as the use of multiple semiotic modes—such as linguistic, visual, spatial, and typographic resources—in constructing meaning. These modes do not function independently but interact to create coherent messages. Cope and Kalantzis (2015) argue that multimodality allows learners to engage with content through multiple channels, thereby enhancing understanding and participation. Similarly, Beauchamp and Kennewell (2018) highlight that visual elements such as images, color, and layout can increase learners' motivation and help clarify complex ideas. In the Indonesian context, several studies (e.g., Sari, 2021; Pratama & Wahyuni, 2022) have

shown that the integration of visual and textual elements in English learning materials can positively influence students' comprehension and engagement, particularly at the senior high school level.

From a social semiotics perspective, meaning is constructed through the interaction of various semiotic resources. Bezemer and Kress (2016) emphasize that the effectiveness of multimodal texts depends on how well these resources are integrated. Jewitt, Bezemer, and O'Halloran (2016) further explain that multimodal analysis enables researchers to examine how visual, linguistic, and spatial elements work together to support meaning-making. However, previous studies in the Indonesian context (e.g., Lestari, 2020; Hidayat, 2021) indicate that although many English textbooks include images and visual layouts, these elements are often used only for decorative purposes rather than as functional components that support deeper understanding. This suggests that there is still a gap between the presence of multimodal elements and their effective use in facilitating meaning-making.

Furthermore, existing studies tend to focus on general multimodal features or students' perceptions, rather than providing a systematic analysis of how specific metafunctions—expression (representational), interaction (interpersonal), and composition (textual)—work together to construct meaning in a textbook. Research applying Kress and van Leeuwen's (2021) visual grammar framework to analyze printed English textbooks, particularly for Grade XII students, remains limited. There is also a lack of in-depth investigation into how these three metafunctions operate simultaneously to enhance comprehension and engagement within a specific instructional material.

The Train of Thought English textbook for Grade XII students represents a multimodal learning resource that integrates text, images, color, and layout. However, the extent to which these multimodal elements function cohesively in terms of expression, interaction, and composition to support meaning-making has not been thoroughly examined. Therefore, this study employs a descriptive qualitative approach and applies Kress and van Leeuwen's (2021) visual grammar framework to analyze how multimodal elements contribute to meaning-making in the textbook.

From the researcher's perspective, multimodality in textbooks is not merely a matter of visual presentation but plays a significant role in shaping how students interpret and understand information. In today's digital era, learners are increasingly exposed to multimodal forms of communication; thus, textbooks must adapt to these developments. Textbooks that effectively integrate multimodal elements can enhance students' comprehension, engagement, and critical thinking. In conclusion, this study aims to provide a comprehensive multimodal analysis of meaning-making—focusing on expression, interaction, and composition—in the Train of Thought English textbook for Grade XII students. The findings are expected to offer valuable insights for teachers, material developers, and researchers in designing and evaluating effective multimodal instructional materials.

METHODOLOGY

This study employed a descriptive qualitative design aimed at analyzing multimodality in the Train of Thought English textbook for Grade XII. The qualitative approach was chosen because it allows for an in-depth interpretation of meaning and the relationship between visual and verbal modes without manipulating the data. As stated by Creswell (2018), qualitative research focuses on understanding social phenomena in their natural context, emphasizing meaning rather than numerical measurement. This approach is therefore appropriate for exploring how multimodal resources function within educational materials. The research adopted qualitative content analysis, which provides a systematic framework for describing and interpreting data through categories and patterns. According to Mayring and Fenzl (2019), this method enables researchers to uncover hidden meanings in texts and visuals by examining how communicative elements interact. In this study, the analysis was based on Kress and van Leeuwen's (2021) theory of visual grammar, which includes three metafunctions: representation, interaction, and composition. These metafunctions were used to examine how images, colors, layout, and textual structures in the textbook create and support meaning.

The data source of this study was the Train of Thought English textbook published by the Indonesian Ministry of Education, Culture, Research, and Technology (Sunengsih et al., 2022). This


Meysha Sandra Kirani Putri^{1*}, Dwi Andriani², Ainur Rohmah³, and Hastuti Retno kuspiyah⁴ textbook was selected because it integrates a variety of multimodal features such as images, diagrams, and digital QR-based learning components, which reflect the current direction of English textbook design in Indonesia. The data were collected through document analysis, involving systematic reading and observation of each unit to identify multimodal elements. The researcher examined how text and visuals appeared together on each page, noting examples of representational meaning (e.g., depiction of characters or actions), interactional meaning (e.g., gaze, gestures, or reader engagement), and compositional meaning (e.g., layout, color, and balance). This process followed the qualitative research cycle described by Miles, Huberman, and Saldaña (2014), where data collection and analysis occur simultaneously and continuously. The data analysis consisted of three steps: condensation, display, and conclusion drawing. During data condensation, visual and verbal elements were categorized according to their metafunctions. In the data display phase, the findings were presented in narrative and tabular form to illustrate how meaning was built through multimodality. Finally, the analysis concluded with an interpretive synthesis of how each metafunction contributed to the construction of meaning in the textbook. Repeated verification was conducted to ensure the consistency and validity of interpretations.



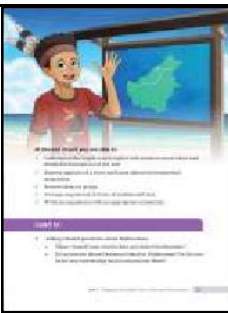


Throughout this process, the researcher acted as the primary instrument of analysis (Flick, 2018), applying both theoretical understanding and interpretive reflection. Moreover, to ensure subjectivity was carefully managed by maintaining transparency in data interpretation and systematically cross-checking the findings against relevant theoretical frameworks. This reflexive approach is essential for bridging theory and data, allowing the analysis not only to describe but also to explain how multimodal resources shape pedagogical meaning. In sum, this methodological approach enabled a comprehensive examination of how textual and visual modes in the Train of Thought textbook interact to construct meaning, with the integration of a descriptive qualitative design, content analysis, and visual grammar theory providing a robust framework for understanding how multimodality supports English language learning in a contemporary educational context.

RESULT AND DISCUSSION






RESULT





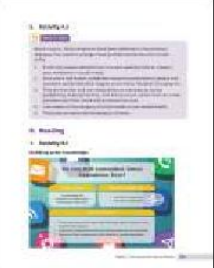
This study analyzed 45 selected pages from the Train of Thought English textbook to examine how multimodal elements, namely representation (R), interaction (I), and composition (C), contribute to meaning making. The pages were selected based on their relevance and the presence of both visual and verbal elements, such as images, text, color, and layout. Each page was analyzed using Kress and van Leeuwen’s visual grammar framework to identify how meaning is constructed through the three metafunctions. The analysis focused on how representation presents ideas and concepts, how interaction engages readers, and how composition organizes information. The detailed findings of each analyzed page are presented in Table 1.

No	Data	Meta function	How to Meta functions Create Meaning
1	 <p>Page xi</p>	R, I, C	The page presents Representation (R) is shown through a road image with labeled signs symbolizing guidance, interaction (I) through an eye-level view creating equality, and composition(C) through central placement and balanced layout, all reinforcing the page’s instructional purpose






2	 <p>Page 1/unit 1</p>	R, I, C	The page presented a student figure examining a board labeled which included photographs, maps, and warning notes. The image involved representation (R) through its depiction of a real- world environmental issue; interaction (I) meta function through the participant’s gaze toward the board, engaging the viewer in an investigative act, and composition (C) through the layout that separated the character from the evidence board, highlighting informational salience.
3	 <p>Page 2/Unit 1</p>	C	The Unit 1 mind map page demonstrated composition through its structured visual layout. Elements such as arrows, text positioning, and the division of activities into four main sections indicated that layout played a central role in conveying meaning.
4	 <p>Page 3 /Unit 1</p>	R, I, C	The page demonstrated multimodality through the integration of visual and verbal elements. The illustration of a boy in traditional clothing and a map of Kalimantan visually supported the unit’s theme (R), while the gesture and gaze of the character created a sense of engagement with the reader (I). The structured layout between the image and the learning objective text reflected a clear composition (C) arrangement.
5	 <p>Page 9/Unit 1</p>	R	The page demonstrated multimodality through the integration of visual and verbal modes. The image showed two characters fishing in a river.
6	 <p>Page 10/Unit 1</p>	R	The image presented multimodality by supporting the narrative text through visual representation. It illustrated Bujang Beji lifting a rock, matching the described event in the story.

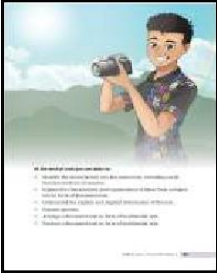




Meysa Sandra Kirani Putri^{1*}, Dwi Andriani², Ainur Rohmah³, and Hastuti Retno kuspiyah⁴

7	 Page 11/Unit 1	R	This page predominantly reflected the representation, as it clearly depicted the dynamic action of Bujang Beji falling. The body posture, facial expression, and visual composition reinforced the narrative in the text, making the image a key element in conveying the story's meaning.
8	 Page 21/Unit 1	I	This page showed the interaction, as it encouraged direct interaction between students, the text, and peers through activity instructions, group work, and the illustrated character addressing the reader.
9	 Page 22/Unit 1	R	This page showed a dominance of the linguistic concept of syllable stress through a clear and simple visual representation.
10	 Page 23/Unit 1	R, I, C	The page primarily reflected the representation (R) through a list of arcade words, supported by interaction (I) elements such as the character gesture and the directive speech bubble. Composition (C) was clearly structured, separating the instruction, illustration, and core vocabulary.
11	 Page 32/Unit 1	R	The page used both text and image to construct a representation of the character Damak. The image showed Damak holding a fish, matching the narrative that described him as a fisherman. This indicated that the visual and verbal modes were used to represent the same event in a complementary way.






12	 <p>Page 38/Unit 1</p>	R	The image and text on the page represented an environmental issue by showing industrial activity in a mountain area and describing its negative effects like floods and landslides.
13	 <p>Page 41/Unit 2</p>	R, I, C	It mainly realized the representation (R) by depicting a young female figure surrounded by various social media icons, symbolizing the online connectivity of youth. Interaction (I) through the indirect gaze and upward-reaching gesture, and composition (C) through the layout that placed the figure at the center as the focus of attention.
14	 <p>Page 42/Unit 2</p>	C	The layout of the page presented a clear division between visual and verbal elements, with the phone illustration on the right and skill descriptions on the left.
15	 <p>Page 43/Unit 2</p>	R	The page strongly conveyed the representational metafunction by depicting visual elements that directly reflected the topic discussed in the accompanying text, enabling readers to easily grasp the intended message.
16	 <p>Page 51/Unit 2</p>	C	The main focus is the systematic arrangement of visual and text information so that the meaning is easy to understand.




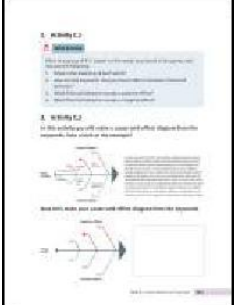

Meysa Sandra Kirani Putri^{1*}, Dwi Andriani², Ainur Rohmah³, and Hastuti Retno kuspiyah⁴

17	 <p>Page 54/Unit 2</p>	R	The page used a conceptual representation to depict online shopping behavior. It showed digital elements like product images, prices, and cart icons to symbolize how consumerism had shifted into the digital space. This supported the accompanying text by illustrating the abstract idea of social media-driven consumption.
18	 <p>Page 55/Unit 2</p>	R	The page employed a conceptual representation, as it visually represented abstract ideas related to digital consumerism. It depicted online shopping features such as price tags, product icons, and a shopping cart symbol to symbolize the shift from traditional to digital purchasing behavior.
19	 <p>Page 61/Unit 2</p>	R	Showing a sad-looking girl surrounded by negative symbols like #BULLY and angry emojis. The visual demonstrates the metafunction of representation.
20	 <p>Page 80/Unit 2</p>	R	The page employed a conceptual representation, as it presented abstract symbols related to social media. The floating icons, such as the heart, hashtag, and other digital symbols, visually represented the online environment that has become integrated into teenagers' daily lives.
21	 <p>Page 94/Unit 3</p>	C	This page applies a centralized Composition where the main theme is placed at the center, surrounded symmetrically by four language skills. This structure organized visual and textual elements cohesively.






<p>22</p>	 <p>Page 95/Unit 3</p>	<p>I</p>	<p>This page displays an illustration of a character smiling and looking directly at the reader while holding a camera, accompanied by text containing learning objectives.</p>
<p>23</p>	 <p>Page 103/Unit 3</p>	<p>R</p>	<p>This page displays a combination of expository text titled "The Hidden Gem of East Nusa Tenggara: Nihiwatu" and an illustration of a surfing figure. The image and text work together to create a visual-verbal multimodality that reinforces the tourism promotion message, reflecting the metafunction of representation.</p>
<p>24</p>	 <p>Page 104/Unit 3</p>	<p>R</p>	<p>This page contains further information about the Nihiwatu tourist destination, accompanied by illustrative images of beach activities. The text and images reflect the metafunction of representation because it presents real-life activities.</p>
<p>25</p>	 <p>Page 107/Unit 3</p>	<p>I</p>	<p>This page shows an illustration of a character inviting students to discuss a trip to Nihiwatu. This image contains the metafunction of interaction because it involves direct interaction between the character and the reader through direct questions in speech balloons.</p>
<p>26</p>	 <p>Page 117/Unit 3</p>	<p>I</p>	<p>The elements show that the illustration invites the reader to chat while explaining the use of asking and giving opinion sentences.</p>





Meysha Sandra Kirani Putri^{1*}, Dwi Andriani², Ainur Rohmah³, and Hastuti Retno kuspiyah⁴

27	 <p>Page 120/Unit 3</p>	R	<p>The page utilized a narrative representation metafunction. It depicted three animated characters engaged in a game. The presence of visible action and emotional expression showed an unfolding event rather than a static concept. This image visually represented social interaction, cooperation, and enjoyment, which reflected the context of the activity described in the textbook</p>
28	 <p>Page 134/Unit 3</p>	R	<p>The image realized a narrative representation metafunction. It showed a character engaged in a travel experience, using a laptop to write, and sharing content online. These three frames depicted a clear sequence of actions (traveling, writing, and posting), which indicated a process rather than a static state. The vector formed by gestures supported the narrative structure, illustrating stages of the holiday-sharing activity described in the text.</p>
29	 <p>Page 139/Unit 4</p>	R,I,C	<p>The page combines three visual metafunctions: representational (R), by depicting modern payment processes such as credit card use and QR code scanning to illustrate technological advancement; interaction (I), through the character's direct gaze and smile that engage viewers and create an approachable tone; and composition (C), by placing the QR codes centrally with supporting elements and bright colors to enhance salience and guide attention to the main concept.</p>
30	 <p>Page 140/Unit 4</p>	C	<p>The layout of each element is arranged and directed with arrows so that the reader understands the purpose of unit 4.</p>
31	 <p>Page 141/Unit 4</p>	I	<p>It shows a character looking at the reader while waving as if greeting them, creating the impression of interaction with the reader. The text element also explains the unit's purpose at the end, indicating that this page falls under the interaction meta function.</p>

32	 <p>Page 152, Unit 4</p>	R, C	The image represents the concept of digital transactions and the future of payment systems. It features a woman smiling with her hand extended toward digital symbols, visually conveying the idea of cashless payments (R). The composition (C), including the QR code and video still, reinforces the connection between the topic and modern financial technology. No interaction is present in the image.
33	 <p>Page 158/Unit 4</p>	R, C	The image represents financial prosperity through a busy city scene with tall buildings and bright advertisements. Its composition directs attention to the subtitle and QR code, guiding viewers to access more information through the linked video.
34	 <p>Page 160/Unit 4</p>	R	The activity page demonstrated multimodality through the integration of verbal and visual modes. The word search grid visually represented the nature of the task, fulfilling the representation meta function.
35	 <p>Page 161/Unit 4</p>	R, C	The page showed a cause-and-effect diagram that illustrated the consequences of using cash or cashless methods. The representation meta function appeared through visual symbols and arrows that conveyed causal relationships. Meanwhile, the composition meta function helped structure the message clearly by organizing ideas from left to right, guiding readers to understand the flow of impact effectively.
36	 <p>Page 164/Unit 4</p>	R, I, C	The image combines verbal (instructions), visual (illustration and thumbnail), and interactive (QR code) elements, making it multimodal. It reflects all three metafunctions: representation (depiction of video making), interaction (gaze from the speaker), and composition (balanced layout guiding attention and meaning).

Meysa Sandra Kirani Putri^{1*}, Dwi Andriani², Ainur Rohmah³, and Hastuti Retno kuspiyah⁴

37	 <p>Page 167/Unit 4</p>	R, I, C	<p>The infographic “Reasons for Shopping” employs multimodal element visuals, text, and spatial layout to convey comparative data between in- store and online shopping preferences. Representation is realized through the statistical data and textual explanations that depict real consumer motivations. Interaction is achieved through the use of color coding and illustrations that engage the viewer. Composition is evident in the side-by-side layout that allows easy comparison, helping the modes work together to construct clear and accessible meaning.</p>
38	 <p>Page 175/Unit 4</p>	R, I, C	<p>The image “Shopping Online: Men vs Women” applies representation through icons and percentages that show shopping preferences. Interaction is built through visual appeal and contrast that attract the viewer. Composition is seen in the balanced layout, clearly dividing male and female data. These elements together help convey the intended comparison effectively.</p>
39	 <p>Page 178/Unit 5</p>	I	<p>The image represents the interaction through a direct gaze, creating a demand that engages the viewer. The focus is on visual interaction, with minimal narrative or compositional meaning.</p>
40	 <p>Page 178/Unit 5</p>	C	<p>Elements (text, color, image) are arranged visually to explain the mind map with a wall-like layout.</p>
41	 <p>Page 179/Unit 5</p>	I	<p>Shows characters greeting readers by waving their hands, accompanied by text containing the final goal of unit 5.</p>

42	 <p>Page 180/Unit 5</p>	R	Shows an explanation of a good lifestyle. Each image symbol represents the explanatory text next to it.
43	 <p>Page 189/Unit 5</p>	R, C	The image functions as a representation of the video linked via the QR code, visually previewing the topic of minimalism and maximalism. Composition also plays a role in guiding the viewer's focus, but there is no significant interaction element present.
44	 <p>Page 201/Unit 5</p>	R	The visual representation on the page illustrated a group of students participating in a debate, which aligned directly with the accompanying instructional text. The characters were shown in dynamic postures with expressive gestures, highlighting the social and collaborative nature of the task.
45	 <p>Page 203/Unit 5</p>	R	The illustration visually represented the concept of frugality by showing piles of coins, reinforcing the topic discussed in the essay prompt. The image served as a metaphor for saving or financial accumulation.

Based on the data presented in Table 1, all three meta functions were identified across the textbook; however, their distribution varies. From the 45 analyzed pages, the **representational meta function appears most frequently**, followed by composition and interaction. Representation is realized through images depicting real-life actions, characters, cultural contexts, and conceptual ideas that complement the accompanying texts. For example, illustrations of environmental issues (Page 1), tourism (Pages 103–104), and daily activities (Pages 10 and 120) visually reinforce linguistic content, enabling learners to better understand abstract or contextual information.

The **interactional meta function** appears through gaze direction, gestures, and speech elements that create engagement between the textbook and the readers. Pages such as Page 21, Page 95, and Page 178 show characters interacting directly with readers through eye contact or inviting gestures, which construct interpersonal meaning and make the learning experience more communicative and engaging. Meanwhile, the **compositional meta function** is reflected in the spatial arrangement of visual and textual elements. Layout, salience, and balance are used to organize information and guide reading flow. This can be observed in mind maps (Page 2 and Page 94), diagrams (Page 161), and infographics (Page 167 and Page 175), where structured visual design helps students follow and understand the content more effectively.

In several cases, all three meta functions are integrated simultaneously (e.g., Pages xi, 23, 139, and

Meysa Sandra Kirani Putri^{1*}, Dwi Andriani², Ainur Rohmah³, and Hastuti Retno kuspiyah⁴ 164), demonstrating a cohesive multimodal design. These pages show how representation, interaction, and composition work together to construct meaning in a more comprehensive way. Overall, the findings indicate that the textbook demonstrates an effective combination of the three meta functions, although **representation remains the most dominant**. This suggests that visual elements are primarily used to illustrate and reinforce textual meaning rather than to maximize interaction or complex compositional design.

DISCUSSION

This study aims to analyze how multimodal elements representation (expression), interaction, and composition to contribute to meaning-making in the *Train of Thought* English textbook. The discussion integrates detailed findings from each data point with the theoretical framework of Kress and van Leeuwen (2021).

First, the **representational meta function (expression)** emerges as the most dominant element across the analyzed data. This is clearly reflected in **Data 5–7 (Pages 9–11)**, **Data 9 (Page 22)**, **Data 11–12 (Pages 32 and 38)**, **Data 15 (Page 43)**, and **Data 17–20 (Pages 54, 55, 61, and 80)**, where images primarily function to depict actions, participants, and real-world or conceptual phenomena. Narrative representations are evident in **Data 6–7 (Pages 10–11)** and **Data 27–28 (Pages 120 and 134)**, where sequences of actions (e.g., Bujang Beji’s story and travel activities) visually illustrate events described in the text. Meanwhile, conceptual representations appear in **Data 17–20**, where abstract ideas such as digital consumerism and social media are visualized through icons and symbols. These findings indicate that meaning-making is largely constructed through visual depictions that transform abstract linguistic content into more concrete and accessible forms. This supports Kress and van Leeuwen’s (2021) argument that representational meaning plays a central role in conveying experiential knowledge. Pedagogically, this dominance suggests that the textbook is designed to facilitate comprehension by linking textual information with visual contexts, which is particularly beneficial for EFL learners.

Second, the **interactional meta function** contributes to meaning-making by establishing a relationship between the text and the reader, although it appears less frequently and more selectively. This can be observed in **Data 8 (Page 21)**, **Data 22 (Page 95)**, **Data 25–26 (Pages 107 and 117)**, and **Data 39–41 (Pages 178–179)**. These pages include elements such as direct gaze, gestures, and speech bubbles that invite reader participation and create interpersonal meaning. For instance, characters looking directly at the reader (Data 22 and 39) or using gestures and dialogue prompts (Data 8 and 25) foster engagement and simulate interaction. This aligns with social semiotic theory, which emphasizes that interactional meaning positions the reader and shapes their response. However, the distribution of this meta function indicates that it is mainly concentrated in activity-based sections rather than consistently integrated across all content. This suggests that while the textbook attempts to engage learners, the use of interactional elements could be expanded to create a more sustained and immersive learning experience.

Third, the **compositional meta function** plays a crucial role in structuring and organizing meaning through layout, salience, and visual hierarchy. This is particularly evident in **Data 3 (Page 2)**, **Data 21 (Page 94)**, **Data 30 (Page 140)**, and **Data 35 (Page 161)**, where mind maps, diagrams, and structured layouts guide students in understanding relationships between ideas. Additionally, **Data 36–38 (Pages 164, 167, and 175)** demonstrate how composition works effectively in infographics by arranging visual and textual elements to

present comparative and statistical information clearly. These findings are consistent with Jewitt et al. (2016), who argue that visual organization is essential in directing attention and facilitating meaning-making. The strong presence of compositional elements indicates that the textbook is systematically designed to support readability and comprehension through clear visual structuring.

Importantly, the analysis also shows that **meaning-making is most effective when the three meta functions operate simultaneously**. This integration is evident in **Data 1 (Page xi)**, **Data 2 (Page 1)**, **Data 10 (Page 23)**, **Data 13 (Page 41)**, **Data 29 (Page 139)**, and **Data 36–38 (Pages 164–175)**. In these cases, representation provides the content, interaction engages the reader, and composition organizes the information. For example, infographics and instructional pages combine visual depiction, reader engagement, and structured layout to create a cohesive and meaningful learning experience. This supports Bezemer and Kress (2016), who argue that meaning is constructed through the interaction of multiple semiotic resources rather than a single mode.

Despite these strengths, the findings also reveal an imbalance in the distribution of meta functions. The dominance of representation, as seen in the majority of data points (e.g., Data 5–7, 11–12, 15, 17–20, 27–28, and 42–45), indicates that the textbook prioritizes the delivery of information through visual illustration. While this approach effectively supports comprehension, it may limit opportunities for deeper interpersonal engagement and interactive learning. Interactional elements, although present, are not consistently integrated, suggesting a potential area for improvement. Therefore, a more balanced integration of representation, interaction, and composition would enhance the overall effectiveness of multimodal learning by not only supporting understanding but also fostering active student engagement.

In conclusion, this study demonstrates that multimodal elements to representation (expression), interaction, and composition for collectively contribute to meaning-making in the *Train of Thought* textbook. Representation facilitates comprehension, interaction fosters engagement, and composition organizes information. However, the imbalance among these meta functions suggests that further optimization is needed to achieve a more holistic and interactive multimodal learning environment.

CONCLUSION

This study aimed to analyze how multimodal elements, specifically representation, interaction, and composition, contribute to meaning making in the *Train of Thought* English textbook for second year high school students. The findings demonstrate that all three metafunctions are present and function collaboratively in constructing meaning, although they are not equally distributed. The representational metafunction emerges as the most dominant, indicating that meaning making is primarily achieved through visual depictions of actions, participants, and concepts that support and clarify the textual content. This dominance highlights the textbook's strong emphasis on facilitating students' comprehension by linking linguistic input with concrete visual contexts.

The interactional metafunction contributes by establishing engagement between the text and the readers through gaze, gestures, and dialogic elements. Although effective in creating interpersonal meaning and encouraging participation, its presence is relatively limited and tends to appear mainly in activity based sections. Meanwhile, the compositional metafunction plays a significant role in organizing information through layout, salience, and visual structure, which helps guide students in understanding relationships between ideas and navigating the

Meysa Sandra Kirani Putri^{1*}, Dwi Andriani², Ainur Rohmah³, and Hastuti Retno kuspiyah⁴
learning material effectively.

Importantly, the study finds that meaning making is most effective when representation, interaction, and composition are integrated simultaneously, as this combination allows visual and verbal elements to complement one another in a coherent way. However, the imbalance in the distribution of these metafunctions, particularly the dominance of representation over interaction, suggests that the multimodal design of the textbook could be further optimized.

In conclusion, the Train of Thought textbook effectively applies multimodal principles to support English language learning, especially in enhancing comprehension through visual representation and structured composition. Nevertheless, increasing the integration of interactional elements would further strengthen student engagement and create a more balanced and holistic multimodal learning experience. These findings provide valuable insights for teachers, material developers, and researchers in designing and evaluating multimodal instructional materials in EFL contexts.

REFERENCES

- Beauchamp, G., & Kennewell, S. (2018). *Interactivity in the classroom and its impact on learning*. Computers & Education, 54(3), 759–766.
- Bezemer, J., & Kress, G. (2016). *Multimodality, learning and communication: A social semiotic frame*. Routledge.
- Cope, B., & Kalantzis, M. (2015). *A pedagogy of multiliteracies: Learning by design*. Palgrave Macmillan.
- Hidayat, A. (2021). Multimodal analysis of English textbooks used in Indonesian senior high schools. *Journal of English Teaching*, 7(2), 85–95.
- Jewitt, C., Bezemer, J., & O'Halloran, K. (2016). *Introducing multimodality*. Routledge.
- Kress, G., & van Leeuwen, T. (2021). *Reading images: The grammar of visual design* (3rd ed.). Routledge.
- Lestari, D. (2020). The use of visual elements in English textbooks: A multimodal perspective. *Jurnal Pendidikan Bahasa Inggris*, 9(1), 45–55.
- Pratama, A., & Wahyuni, S. (2022). The use of multimodal elements in English textbooks for secondary school students. *Journal of English Language Teaching*, 11(1), 45–53.
- Sari, D. (2021). The role of multimodal texts in improving students' reading comprehension. *Jurnal Pendidikan Bahasa Inggris*, 10(2), 123–130.