

CHAPTER 2

LITERATURE REVIEW

This chapter gives detailed information about the theories from experts related to the research. The literature review explains about the theoretical framework being used in the research. Furthermore, the relevant previous study is attached to give more relevant data and information regarding this research.

A. Theoretical Framework

This study uses an integrated theoretical framework that combines Piaget's cognitive development theory, the revised Bloom's taxonomy along with types of knowledge, the concept of cognitive representation, and two analytical approaches—SFL transitivity for the verbal mode and MDA for the visual mode. These theories are integrated because each explains different aspects of meaning construction in textbooks: children's thinking capacity, forms of mental processes and knowledge represented, patterns of information presentation, and how language and images work together to construct meaning multimodally.

1. Cognitive Development Theory

Jean Piaget's theory of cognitive development provides a conceptual foundation for understanding how children build knowledge and develop ways of thinking as they grow older. Piaget asserted that intellectual development does not occur linearly, but through qualitative changes in the child's mental structure (Piaget, 1952). These changes stem from the active interaction between the child and their environment, so the learning process is always related to how the child organizes, interprets, and balances new experiences.

Piaget explained that the basic structure of cognitive development lies in schemas, which are mental frameworks that organize knowledge. Sanghvi (2020) explained that schemas develop from infancy through sensorimotor actions such as grasping, staring, or repeating movements, which gradually evolve into more stable patterns of mental representation. When faced with new information, a child engages in assimilation, which involves

incorporating that experience into existing schemas; whereas accommodation occurs when new information requires changes to old schemas in order for understanding to function properly. These two mechanisms operate within the dynamics of adaptation, which is a child's effort to achieve a balance between existing knowledge structures and new experiences, a state referred to by Piaget as equilibration. This mechanism is consistently described by various researchers, including Rabindran & Madanagopal (2020), and is elaborated in detail in Saul McLeod's discussion (2018), which emphasizes that new development occurs when a child successfully balances conflicts between prior understanding and new information.

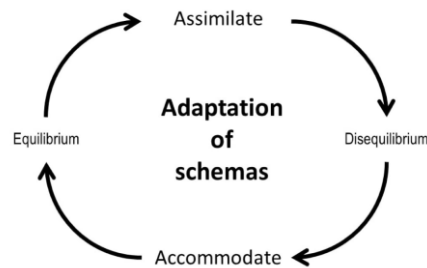


Figure 1. Scheme Adaptation Cycle: Assimilation, Accommodation, and Equilibration.
Source: McLeod (2018).

This foundation of cognitive development cannot be separated from the sensorimotor stage. In the book *The Origins of Intelligence in Children* by Piaget, it is detailed how physical actions in the early years of life form the basis of thinking. Piaget demonstrated that early reflexes develop into repetitive action patterns (primary and secondary circular reactions), then evolve into goal-directed exploration, and at the final sensorimotor stage result in mental representation, which is the ability to maintain an internal image of an object without needing direct contact with it. This stage becomes the foundation for a child's symbolic abilities, which later becomes important in the preoperational period.

In the preoperational stage (around ages 2–7), children begin to demonstrate symbolic thinking, allowing the use of language, images, and objects as representations of reality. However, their thinking is still intuitive and tied to perception. Piaget (1952) emphasized that children at this stage

experience centration, which is the tendency to focus attention on one prominent visual feature, as well as egocentrism, which is the difficulty in understanding other people's perspectives. McLeod (2018) adds that children's thinking is still perceptually bound, so reasoning relies heavily on physical appearances rather than underlying concepts. Rohmah et al. (2022) also indicate that preoperational children do not yet understand conservation, reversibility, and logical relationships between elements, so learning materials for this age group should ideally be explicit, concrete, and not require high-level abstraction. The symbolic abilities that develop at this stage are an important prerequisite for understanding texts and images, but their intuitive limitations demand simple and straightforward visual and verbal representations.

Entering the concrete operational stage (around ages 7–11), children show significant changes in thinking ability. They begin to engage in logical reasoning, but only as long as the reasoning is related to concrete objects or real situations (Piaget, 1964). Children at this stage are able to perform classification and seriation, understand reversibility, and begin to show decentration, that is, attention to more than one aspect of a situation. Rabindran & Madanagopal (2020) emphasize that this ability is a key feature of the shift from intuitive thinking to concrete logical thinking. McLeod (2018) adds that children's understanding of simple cause-and-effect relationships begins to improve at this stage, although full abstraction has not yet developed. Educational research by Pakpahan & Saragih (2022) shows that elementary school children understand concepts more quickly when presented through real objects, concrete illustrations, or direct experiences, because their reasoning is still very much tied to empirical context.

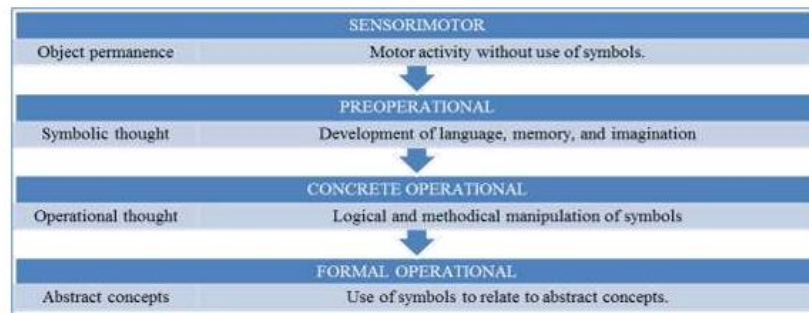


Figure 2. Stages of Cognitive Development According to Piaget

Source: Rabindran & Madanagopal (2020).

Children aged six to seven are in an interesting transitional stage between the end of the preoperational stage and the beginning of the concrete operational stage. At this stage, children already have fairly stable mental representation abilities, but they still need visual support and concrete situations to understand logical relationships (McLeod, 2018). They begin to be able to group objects based on simple categories, follow action-based instructions, and discover patterns or relationships between objects. However, they are still easily distracted by dominant visual features and often require explicit examples to connect text or visuals with everyday experiences (Huwaida et al., 2023). Kholiq (2020) emphasizes that early elementary school children derive meaning primarily through concrete experiences, while abstract concepts can only be understood if accompanied by clear symbolic representations.

These developmental characteristics have direct implications for how children understand verbal and visual representations in textbooks. On the verbal side, children find it easier to process language related to actions and real objects, as explained by Piaget, who stated that children's logical thinking is still tied to concrete experiences (Piaget, 1964). Clause structures that describe direct activities or simple identification relationships tend to be easier to understand, as they align with the way children build meaning through actions and observations. On the visual side, illustrations that use real objects, contrasting colors, clear proportions, and simple compositions support children's tendency to focus on one important element before

expanding their focus. Visual representations that are too dense or too abstract tend to disrupt the meaning-making process because children are still heavily influenced by the physical appearance of objects.

Overall, Piaget's theory provides a strong framework for understanding how early elementary school children process symbols, build meaningful relationships, and integrate verbal and visual information. Understanding this developmental stage allows for the analysis of representations in textbooks to be carried out with consideration of the cognitive characteristics of children who are in the transitional phase toward concrete reasoning. Thus, Piaget's theory serves as a foundation for interpreting the forms, characteristics, and structures of representations in Rise and Shine grade 1, as well as how these representations relate to the way children construct meaning based on their cognitive development.

2. The Process of Cognitive Development

Understanding the process of cognitive development not only includes the stages of thinking abilities but also how knowledge is processed, organized, and utilized in learning activities. A framework widely used to explain these two aspects in an integrated manner is Bloom's Revised Taxonomy developed by Anderson & Krathwohl (2001). This revision not only updates the cognitive process categories from Bloom's (1956) taxonomy but also adds a knowledge dimension to address the limitations of the old framework, which did not systematically classify types of knowledge. The result is a two-dimensional taxonomy that combines mental processes with the structure of knowledge, providing a more comprehensive picture of how learners study and understand information.

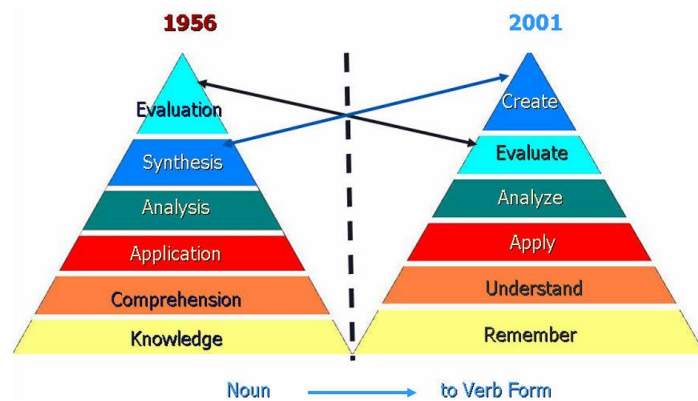


Figure 3. Bloom's Revised Taxonomy (Anderson & Krathwohl, 2001), adapted from Wilson (2016, p. 4)

In the first dimension, cognitive processes are represented through six categories arranged by complexity: remembering, understanding, applying, analyzing, evaluating, and creating. The shift from noun-based categories (such as “knowledge” or “comprehension”) to verbs reflects an emphasis that learning is an active mental activity. Remembering refers to the ability to recall basic information; understanding includes the ability to interpret and explain; applying indicates the use of knowledge in new contexts; analyzing involves breaking information into parts; evaluating includes assessment based on certain criteria; and creating requires the ability to design, organize, or produce new ideas. This hierarchy is shown in Figure 3, which displays a comparison between the original Bloom's taxonomy and its revision, and highlights the placement of “creating” as the most complex cognitive process.

The second dimension, the knowledge dimension, describes the types of knowledge processed in learning activities. Anderson & Krathwohl (2001) group knowledge into four categories: factual knowledge, conceptual knowledge, procedural knowledge, and metacognitive knowledge. Factual knowledge includes basic elements such as terms, symbols, and concrete information that form the initial foundation of learning. In early elementary school, this form of knowledge is usually represented through labels, basic vocabulary, or pictures showing familiar objects. Conceptual knowledge refers to relationships between concepts, such as categories, classifications, or part-whole patterns. This knowledge is important for early learners as they

begin to understand the grouping or characteristics of objects through text or visuals. Procedural knowledge includes the steps or procedures for carrying out an action. In 1st grade material, this appears through simple instructions that are realized through material processes or illustrations of sequential actions. Metacognitive knowledge, although still very basic at this age, includes awareness of one's own thinking processes, such as checking answers or making simple decisions.

This two-dimensional taxonomy provides a flexible framework for understanding how children aged 6–7 process information. At the concrete operational stage, lower-level cognitive processes—remembering, understanding, and applying—tend to become more dominant. Children are better able to understand material when the information provided is concrete, structured, and related to direct experiences. Therefore, the emergence of cognitive representations at the LOTS (lower-order thinking skills) level in early grade textbooks reflects a match with developmental capacity, rather than an indication of weak material. Similarly, factual and simple conceptual knowledge appears more frequently compared to complex procedural knowledge or abstract metacognitive knowledge.

In this descriptive study, Bloom's taxonomy and knowledge dimensions are used to map the types of mental processes and types of knowledge represented through verbal and visual modes. Grammatical structures such as labelling or identifying are often associated with remembering and understanding; material processes show applying; while part-whole visualization or object grouping can indicate conceptual knowledge. Thus, this framework allows for a systematic identification of the cognitive demands built into the teaching material without making an assessment of the design's effectiveness. The integration between cognitive processes and types of knowledge provides an important foundation for interpreting patterns of cognitive representation expressed through language and visuals in the Rise and Shine grade 1 English textbook.

3. Cognitive Representations

Cognitive representation is a way of organizing knowledge in teaching materials so that readers can understand it more clearly. This concept builds on the idea that knowledge must be arranged in a certain form so that readers, especially children, can capture its meaning. In other words, cognitive representations serve as "patterns of meaning presentation" that help learners see the relationship between one piece of information and another. The pattern includes the way concepts are introduced, connected, and arranged so that they form a structure that can be logically followed. In the context of textbooks, cognitive representation becomes a bridge between the content of the material and the reader's way of thinking at a certain stage of development.

In early grade textbooks, cognitive representation is usually realized through a combination of text and visuals. Verbal text provides an explanation, naming, or description, while illustrations show real shapes, actions, or relationships between the elements depicted. The two work together to make it easier for children to understand the meaning presented on a particular page. For novice readers, clear images are often the first entry point for understanding, while text serves to reinforce or complement the meaning constructed by visuals. In this study, the concept of cognitive representation is used to identify the forms of knowledge presentation in the Rise and Shine grade 1 textbook. Thus, cognitive representation becomes an important link between cognitive development theory and multimodal approaches, and also provides a conceptual basis for understanding how verbal and visual texts work together in building student understanding.

4. Systemic Functional Linguistics

Systemic Functional Linguistics (SFL) views language as a resource for constructing meaning through three main metafunctions, namely ideational, interpersonal, and textual (Yanto & Hikmah, 2023). In this study, attention is mainly directed to the ideational metafunction, specifically experiential meaning, which explains how language represents experiences, actions, feelings, and relationships in the world. Halliday & Matthiessen (2004a) explain that experience is realized through a process-participant-

circumstance structure in clauses. With this framework, each clause can be seen as a single unit of linguistically constructed experience. The transitivity analysis then allows the researcher to map how clauses in textbooks form representations of certain experiences for early elementary school children.

Transitivity consists of six types of processes, namely material, mental, relational, behavioral, verbal, and existential. Each type of process builds a different domain of experience and presents a specific way of looking at the world. Material processes describe actions or events that are concrete and observable, such as run, open, or pick, so they are very close to the way 6–7-year-olds understand the world through visible actions. Mental processes include perception, affect, and cognition, such as see, like, or know, which reflect inner experiences and help children relate texts to their perceptions or emotions. The relational processes function to classify or identify something through verbs such as is or has, and play an important role in the formation of basic concepts, such as colors, shapes, or categories of objects. Behavioral processes are somewhere between material and mental processes, such as laugh, cry, or watch, while verbal processes, such as say or ask, construct speech actions that often appear in children's texts that feature dialogue. Finally, existential processes mark the existence of something through patterns such as there is/are, which are often used when introducing new objects or situations in a lesson.



Figure 4. Process Types in the Transitivity System (adapted from Halliday & Matthiessen, 2004, p. 172).

In addition to the processes, transitivity also includes the participants who play a role in the experience. Halliday & Matthiessen (2004a) emphasized that each type of process has a distinctive participant configuration, such as Actor/Goal for material processes, Senser/Phenomenon for mental processes, Carrier/Attribute or Token/Value for relational processes, and Sayer/Receiver for verbal processes. This configuration shows who performed the action, who experienced it, and what objects were involved in an event. For grade 1, the clarity of the participants is very important because children at the concrete operational stage find it easier to understand concrete entities with stable and easily recognizable roles. Therefore, the analysis of participants in textbook clauses helps to describe how actions and events are structured to be accessible to the child's way of thinking.

Transitivity also includes elements of circumstances, such as place, time, manner, purpose, or accompaniment. The element of circumstance is grammatically optional, but it is very helpful for 6–7-year-olds to relate events to a concrete context, such as "in the park", "with my friend", or "in the morning". Halliday & Matthiessen (2004a) emphasize that circumstances enrich the structure of experience and clarify the relationships between events in the text. In the context of the Rise and Shine textbook for grade 1, the existence of circumstances can show the extent to which the verbal content provides clear markers of space and time so as to support the child's need for a real context. Thus, mapping of circumstantial elements is an important part of the transitivity analysis in this study.

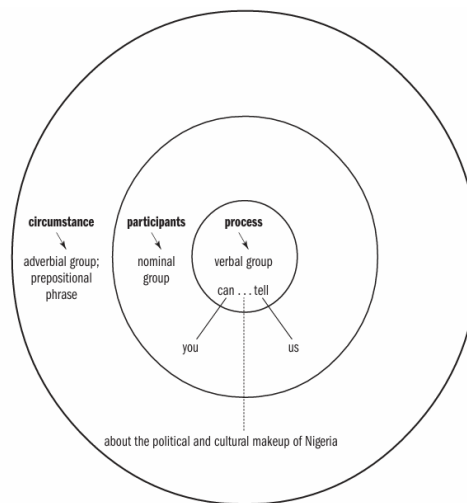


Figure 5. *Experiential Structure of the Clause* (adapted from Halliday & Matthiessen, 2004, p. 176).

The transitivity framework becomes relevant when it is linked to the theory of cognitive development and the cognitive representation framework used in this study. The dominance of certain types of processes, participant choices, and patterns of action depictions can provide clues about the forms of cognitive representation that emerge, such as definitions, classifications, sequences of actions, or causal relationships. For grade 1 children who are still in the concrete operational stage, representations that are too abstract, such as very nominal clauses or high-level mental processes, are potentially less in line with the way they process information. In this study, transitivity analysis is used to describe how verbal language works with visual elements in building the learning experience offered by textbooks, without intending to assess the quality of the language. In that way, SFL provides a linguistic basis for understanding the patterns of cognitive representation present in the Rise and Shine textbook for grade 1.

5. Multimodal Discourse Analysis

Multimodal Discourse Analysis (MDA) is based on the view that meaning in communication is not only formed through language, but through the combination of various semiotic modes such as images, colors, typography, layout, and gestures (Jewitt et al., 2016). Each mode has a meaning potential, that is, the semiotic capacity to convey certain meanings

according to the social context and communicative purpose. In the context of children's education, this framework is particularly relevant because textbooks for early readers are almost always designed as multimodal texts. Students' understanding of the content of the lesson depends not only on the words, but also on the accompanying visuals, including the way images, colors, and layouts are directed to guide attention. Therefore, the analysis of textbooks needs to consider how these various modes interact with each other in building meaning for early readers.

The main theoretical framework in visual analysis is the Grammar of Visual Design (GVD) developed by Kress & Van Leeuwen (2006) by adapting Halliday's metafunctions into the visual realm. They identified three types of visual meanings, namely representational, interactional, and compositional meanings. Representational meaning describes experience through narrative structures, such as actions or events, as well as conceptual structures, such as classification and part-whole relationships. Interactional meaning deals with the relationship between the image and the viewer through the direction of view, social distance, angle, and visual attitude. Meanwhile, compositional meaning regulates the layout of visual elements through information value, salience, and framing. In this dimension, color, font size, and placement of visual elements help determine which information seems more important to the reader. A study of basic color terms by Berlin & Kay (1969) argue that certain basic color terms (e.g., red, yellow, blue, green) are widely attested and perceptually salient across languages. In the context of visual design, using such basic, contrasting colors can strengthen visual salience in children's book illustrations.

The development of MDA further gave birth to the Systemic Functional-Multimodal Discourse Analysis (SF-MDA) approach which emphasizes the functional relationship between language and other modes in the process of meaning. The basic principle of SF-MDA is intersemiotic complementarity, namely the way two or more modes work together to produce meanings that cannot be obtained when each mode is read separately

(O'Halloran, 2008; Royce, 2007). For example, material processes in text—verbs such as run, jump, or draw—can be emphasized through narrative visuals that display the same actions. In contrast, relational processes, such as the identification of "This is a cat", are often supported by conceptual visuals that display objects statically and clearly. In this way, verbal and visual texts form an integrated and complementary representation of knowledge. A similar approach was used by Andriani et al. (2024) in analyzing specific chapters of ELT Indonesian textbooks, which examined how texts and illustrations work together to build the meaning of religious values. This approach is in line with the descriptive nature of the research because it allows mapping of intermodal relationships without having to conclude an assessment of the quality of book design.

The concept of intersemiotic complementarity becomes particularly relevant when it is associated with the cognitive development of children aged 6–7 years. At the concrete operational stage according to Piaget, children begin to be able to do logical reasoning, but still rely heavily on concrete and visual representations. Visuals in textbooks serve as a cognitive scaffold, which helps children interpret verbal texts. Illustrations with easily recognizable objects, vivid colors, proportional sizes, and explicit sequences of actions can help children overcome their tendencies toward centration and being perceptually bound. The use of prominent basic colors can also make it easier for the child to group and recognize objects, in line with their perceptual tendencies at this stage. Visualizations designed in such a way facilitate the focus of attention and interpretation for the early reader, so that the meaning constructed through language is more easily connected to everyday experience.

The MDA approach can also be directly linked to the Bloom's Revised Taxonomy and the knowledge dimension. Visuals that display objects with simple labels, for example, support the cognitive processes of remembering and understanding, while action sequence illustrations can encourage the process of applying. Visualization of patterns, classifications, or part-whole

relationships is closely related to conceptual knowledge, while procedural images that show step-by-step can represent procedural knowledge. Thus, the integration of verbal and visual elements allows the formation of knowledge representations that are in harmony with the stages of children's cognitive development. In the context of the analysis of the Rise and Shine grade 1 textbook, MDA provides a framework to explore how text and illustration are combined as a pedagogical unit and how these combinations form a cognitive representation for the initial reader without providing an evaluative assessment of the book being studied.

6. Framework Integration

The theoretical framework in this study combines several perspectives that work at different levels of analysis, namely cognitive development, thought processes, knowledge structure, language representation, and visual representation. Integration is necessary because the cognitive representations in textbooks do not arise from a single source, but rather from the relationship between the child's abilities, the forms of knowledge introduced, and the way language and visuals construct the learning experience. Within this framework, Piaget's theory of cognitive development is a starting point because it provides an overview of the meaning-making capacity of grade 1 students. Children aged six to seven years are in a transition phase to the concrete operational stage, when understanding is still highly dependent on concrete and observable representations. They begin to be able to recognize simple patterns, group objects, and follow basic cause-and-effect relationships. By understanding the limits and potential of these developments, the analysis can identify the forms of representation that the early reader is most likely to understand, either through text or images.

Bloom's Revised Taxonomy complements Piaget's perspective by mapping out the mental processes that matter might demand. At the grade 1 level, the process of remembering and understanding is generally more dominant because it is in harmony with basic recognition and comprehension needs. The process of applying or analyzing can appear in action instructions

or activities that involve observing simple patterns. In this study, the framework is used descriptively to observe variations in cognitive demands in textbooks, without assessing whether these levels are "adequate" or not. The dimension of knowledge presented by Anderson & Krathwohl (2001) clarifies the form of information that books offer, ranging from factual knowledge in vocabulary and object labels, to conceptual knowledge in categories and relationships, to procedural knowledge in simple steps that students need to follow, to metacognitive knowledge that minimally represented at this level. These four forms help explain how knowledge structures are built in the material and how novice readers can access it according to their cognitive capacity.

Meanwhile, SFL provides a linguistic basis for examining how experiences are represented through language. The transitivity system allows the analysis of the types of processes used, such as material processes that express concrete actions, relational processes that establish identification or classification, and mental processes that describe simple perceptions. The choice of these process types helps to uncover patterns of representation of meaning formed through verbal clauses, especially those relevant to the child's way of thinking. Multimodal analysis through MDA complements the linguistic dimension by examining visual roles, including images, layouts, typography, and color. Through representational, interactional, and compositional structures, visuals in textbooks can display actions, direct the reader's attention, or organize information into a part-whole relationship. Within this framework, color is seen as an important semiotic resource; the choice of prominent basic colors, as discussed in the study of basic color terms by Berlin & Kay (1969), can help children recognize and group objects more easily. The SF-MDA approach allows the relationship between modes to be read through the principle of intersemiosis, which is how text and images work together to produce a complete representation of knowledge without making normative judgments about which patterns are more or less effective.

Overall, the integration of these theories results in a thorough analytical framework. Piaget's theory describes the cognitive capacity of the reader, Bloom's Taxonomy maps the mental processes that may be demanded, while the dimension of knowledge describes the structure of the information presented. SFL explains how language organizes the meaning of experience, while MDA shows how visual elements complement it in a multimodal whole. By bringing all these perspectives together, this study seeks to describe how the Rise and Shine grade 1 textbook builds cognitive representation in a multimodal manner. The focus of the analysis lies in mapping the patterns that emerge and their alignment with the characteristics of the children's development, not on assessing the quality or effectiveness of the material. Thus, this theoretical framework supports the descriptive research goal, which is to describe how textbooks construct learning experiences for early learners.

B. Study of Relevant Research

Research on textbooks, particularly those that examine cognitive representations and the demands of thinking, shows that textbooks function not only as a source of material, but also as an instrument that shapes the way students organize knowledge. Omar (2020), for example, analyzed cognitive representations in grade 9 science textbooks and found that concepts are presented more through conceptual knowledge than procedural, with concept integration that tends to be explicit. This study confirms that the way concepts are introduced (e.g., through analogies, linguistic contexts, and relationships between concepts) has an effect on the likelihood of students' misconceptions. Similar findings emerged in studies of cognitive level evaluation of questions and assignments in textbooks, such as Putri et al. (2023) in SMK English books, and Ulum & Taşkaya (2019) in Turkish language textbooks for primary school, both of which showed the dominance of the level of remembering and understanding. However, most of these studies focused on the middle and upper levels, so there is still limited research that specifically maps how cognitive

representations are constructed in early grade textbooks, especially in the context of EFL

A number of other studies used Bloom's Revised Taxonomy to map cognitive demands in textbooks in different countries. Sainyakit (2023) analyzed the book *Grow with English* for grade 6 of elementary school and found that activities are still dominated by low cognitive processes (remembering, understanding), with limited portions for higher levels. Hong (2024) examined assignments in grade 7 English books in China and showed that Higher Order Thinking Skills (HOTS) only appeared in a small number of activities. At the high school level, Xie (2024) also reported that cognitive demands in EFL books tend to not consistently encourage students' analytical and evaluative abilities. In Indonesia, Rinjaya & Halimi (2022) and Sabir et al. (2024) used Bloom's taxonomy to evaluate textbooks and teacher questions, and both affirmed that high cognitive levels are still rarely systematically facilitated. In general, these studies show a pattern of dominance of LOTS in various contexts, but not many have explicitly linked it to the cognitive development characteristics of children aged 6–7 years.

On the other hand, multimodal studies of textbooks provide an overview of how language and images work together in constructing meaning. Hidayat et al. (2023a) analyzed the images in the *Get Smart Plus* book with the framework of Grammar of Visual Design and the transitivity system, and found representational, interactional, and compositional patterns that complement each other in verbal text in aiding student comprehension. Rafelina & Hermawan (2022) shows that the multimodal texts in the EFL book make extensive use of relational and verbal processes to build ideational meanings, with visual support that clarifies the identity and attributes of objects. Other studies, such as Fristiara (2018), Ghouschi (2021), and Jamilah et al. (2024), also confirm that the images in the EYL/EFL book are not just decorations, but rather important modes that direct attention and organize information. However, most of these studies stop at text-image relationship mapping and have not examined how

these multimodal patterns relate to certain types of cognitive representations or thinking demands.

Further development of the Systemic Functional–Multimodal Discourse Analysis (SF-MDA) framework is seen in studies that combine SFL with visual analysis. Xiangping & Xiaodong (2022) analyzed the discourse of PEP's high school English books and showed how the choice of processes, participants, and visual elements shaped the representation of experiences and social roles in the classroom. Yanuar et al. (2021) specifically highlighted intersemiosis in the ELT book, i.e., how text and images complement each other in constructing meaning, and concluded that effective intersemiosis can improve students' multimodal literacy. Stec (2021) and Sugianto et al. (2021) reviewed ELT materials for young learners and found that a combination of layout, color, and ideational structure has the potential to support cognitive development if designed consciously. These studies show that SF-MDA provides a powerful analytical tool for reading intermodal relationships in textbooks, but the focus is still rarely directed on conformance to the early stages of cognitive development in elementary school.

In addition, some studies have begun to combine cognitive and multimodal perspectives more explicitly. Soe (2024) through a systematic review of ELT book evaluations shows that most studies still use checklists and surface analysis, while studies that integrate the Bloom, SFL, and MDA frameworks are still limited. Lizasoain & Vargas (2023a) highlight how lexical complexity in textbooks can build everyday knowledge gradually, while Song and Perry (2023) show that intersemiotic complementarity in communicative course books plays an important role in the development of multimodal literacy. In Indonesia, Hidayat et al. (2023a) and Hafiz et al. (2024) indicate that a multimodal approach has great potential to examine how textbooks support or inhibit students' thinking processes. However, to date, very few studies have specifically analyzed elementary grade 1 English books by combining SF-MDA, Piaget, Bloom, and knowledge types frameworks to map cognitive representations in detail.

Based on the research map, the position of this study becomes clearer. First, studies on cognitive representation in textbooks have been carried out in the field of science and at the secondary level, but have not touched on the specifics of early grade EFL books with the Indonesian context. Second, research using Bloom's Revised Taxonomy generally focuses on analyzing questions or assignments, rather than on how text and images together build knowledge structures. Third, multimodal studies and SF-MDA have shown the importance of verbal–visual connections, but it is still rare to link them to the characteristics of cognitive development in children aged 6–7 years. Fourth, the thorough integration between SFL, MDA, Piaget, and Bloom in analyzing textbooks for early readers is still very limited, especially for international books used in Indonesian private elementary schools. Thus, this study seeks to fill this gap by describing how Rise and Shine Grade 1 builds cognitive representation in a multimodal manner, as well as the extent to which the patterns that emerge are in harmony with the cognitive capacity of grade 1 students.