



OPEN ACCESS

EDITED BY
Vassilis Martiadis,
Department of Mental Health, Italy

REVIEWED BY
Pauline Jones,
University of Wollongong, Australia
Fabiola Raffone,
Department of Mental Health, Italy

*CORRESPONDENCE
Deborah L. Nichols
✉ deborahnichols@purdue.edu

RECEIVED 31 March 2025
ACCEPTED 24 June 2025
PUBLISHED 22 August 2025

CITATION
Nichols DL, Neal T and Hand B (2025) A
hierarchical typology of multimodal scientific
writing in early primary grades.
Front. Commun. 10:1603737.
doi: 10.3389/fcomm.2025.1603737

COPYRIGHT
© 2025 Nichols, Neal and Hand. This is an
open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](#). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic practice.
No use, distribution or reproduction is
permitted which does not comply with these
terms.

A hierarchical typology of multimodal scientific writing in early primary grades

Deborah L. Nichols ^{1*}, Ted Neal ² and Brian Hand ²

¹Human Development and Family Science, Purdue University, West Lafayette, IN, United States,
²Department of Teaching and Learning, College of Education, University of Iowa, Iowa City, IA,
United States

Introduction: Scientific literacy in early education depends on both content knowledge and young learners' ability to organize and express scientific ideas across linguistic and visual-graphical modes. This integrative ability, known as multimodal competency, enables children to construct and communicate scientific understanding by coordinating text structure, language, and visual representation.

Methods: In this study, 1,705 science writing samples from 1,008 students in kindergarten through 2nd grade were analyzed using a two-stage quantitative approach. First, a content analysis was conducted in which each sample was systematically coded for text structure, linguistic and visual-graphical modes, and scientific content. These coded features were then subjected to latent class analysis (LCA) to identify distinct typologies of multimodal text production.

Results: LCA identified six typologies of multimodal composition, ranging from simple unlabeled visual texts to strategically integrated complementary compositions. These typologies capture meaningful variation in how students organize and communicate scientific ideas, reflecting differences in metacognitive strategy use, modality coordination, and sensitivity to the communicative functions of linguistic and visual modes.

Discussion: Findings indicate that young children's multimodal texts act as both communicative and epistemic tools, supporting scientific meaning-making. Rather than requiring explicit instruction, just-in-time scaffolds may help children make strategic representational choices, reinforcing the role of multimodal composition in early science learning.

KEYWORDS

multimodal writing, primary science education, young learners, scientific text production, linguistic and visual-graphical modes, epistemic tools, generative learning, metacognition

Introduction

Developing a well-organized knowledge base is fundamental for effective learning in science. Experts, as opposed to novices, possess a broader and more systematically structured knowledge base, enabling them to categorize and integrate new information efficiently (Bilalić et al., 2010; Bransford et al., 2000; Feltovich et al., 2006). The process of organizing and integrating knowledge is influenced by, and concurrently influences, a range of cognitive processes, including metacognition, the strategic selection and integration of information, and the utilization of a semiotic proficiency to incorporate different representational modes. These cognitive abilities allow learners to dynamically structure new knowledge, critically evaluate their learning strategies, and leverage the unique affordances of different modes to enhance scientific understanding and expression (Bransford et al., 2000; Chi et al., 1981). However, as Brod (2021) argues,

the ability to effectively engage with generative learning strategies particularly those that involve using language and other modes as tools for constructing understanding is constrained by cognitive development. Younger learners may not yet possess the metacognitive skills required to actively structure and integrate knowledge in the same way as older learners. [Fiorella and Mayer \(2015\)](#) propose that generative learning involves eight core strategies including summarizing, mapping, and self-explaining, all of which are necessary for scientific meaning-making but require varying levels of cognitive maturity. This raises important questions about the developmental trajectory of multimodal competency and, specifically, how young children begin to use language and visual representations as epistemic tools in scientific writing.

Writing as an epistemic tool in early science learning

Language is more than a medium for communication; it is also an epistemic tool that allows learners to construct, refine, and extend their understanding ([Prain and Hand, 2016](#)). Writing serves a dual function. It both externalizes knowledge and supports generative learning by requiring students to integrate, organize, and elaborate on scientific concepts. This epistemic function of writing is well established in disciplinary literacies, where scientific texts represent more than records of knowledge; that is, they serve as mechanisms for knowledge construction and transformation ([Norris and Phillips, 2003](#); [Kress, 2010](#); [Mayer, 2020](#)). For young learners, the ability to use writing as an epistemic tool is still emerging. Generative learning theory ([Wittrock, 1992](#)) suggests that making meaningful connections between new and existing knowledge requires storing, transforming, and linking information across different cognitive and representational modes. In scientific writing, this means that learners must produce coherent linguistic structures and coordinate multiple modes of representation, including visual-graphical elements ([Kress, 2010](#)). The extent to which early learners develop multimodal competency, defined as the ability to selectively integrate linguistic and graphical elements into cohesive scientific explanations, remains an open question.

The role of multimodal competency in science communication

Within the scientific community, writing is a primary medium for knowledge generation, dissemination, and conceptual refinement ([Norris and Phillips, 2003](#); [Yore et al., 2003](#)). The National Research Council [[National Research Council \(NRC\), 2017](#)] explicitly recognizes the centrality of multimodal communication in science, stating that: “A major practice of science is thus the communication of ideas and the results of inquiry—orally, in writing, with the use of tables, diagrams, graphs, and equations...” (p. 53). For young learners, developing the ability to navigate and coordinate these multiple semiotic systems is essential for early scientific literacy ([Fang, 2006](#)). Our understanding of how children develop multimodal competency, particularly how they shift between and eventually integrate communicative and

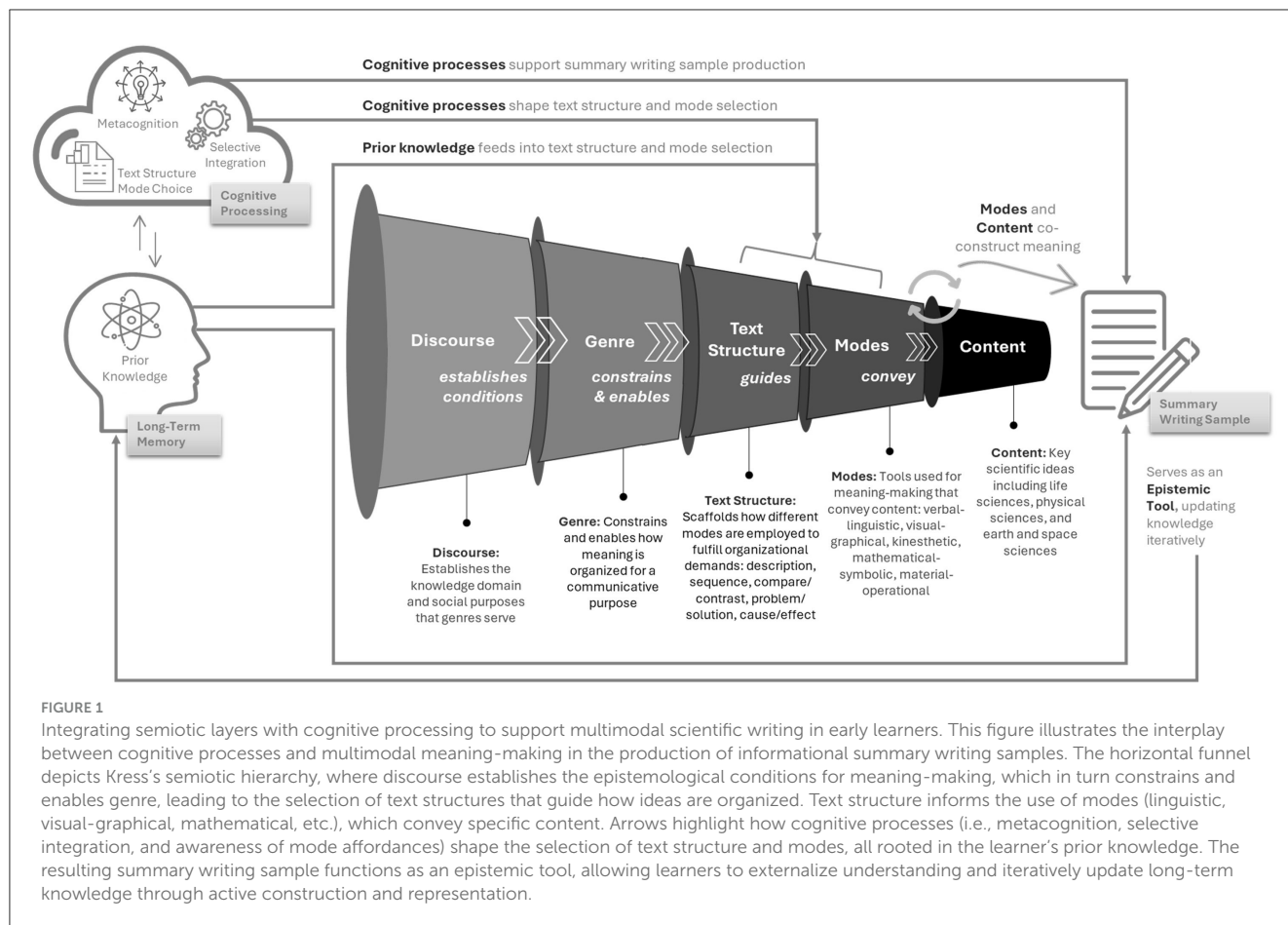
epistemic uses of language and representation, remains limited. Rather than representing a strictly sequential transition, this process may involve recursive and context-dependent shifts in how students use modes for expressing and constructing knowledge. While prior research has primarily focused on the benefits of multimodal representation for knowledge retention and conceptual understanding, less is known about how different modal forms function in the early generation of scientific meaning ([Palincsar et al., 2001](#); [Prain and Hand, 2016](#); [Ainsworth et al., 2020](#)). Future research that incorporates cognitive, semiotic, and sociocultural dimensions may offer a fuller picture of how these representational competencies emerge and evolve in classroom settings.

Study purpose and research questions

Using a content analytic approach, this study examines how young learners structure their scientific knowledge through text production. Unlike prior research that focuses on modal affordances in isolation, we investigate how text structure, mode use, and content interact to shape early multimodal competency in science communication. Specifically, we analyze variations in discourse (i.e., how children construct and organize scientific explanations); genre (i.e., the types of scientific texts produced and their rhetorical structures); text structure (i.e., how information is hierarchically arranged in writing); mode use (i.e., the integration of linguistic, graphical, and visual elements); and content (i.e., the conceptual depth and coherence of scientific representations). By identifying patterns in how children coordinate these elements, it is possible to identify the cognitive and strategic mechanisms underlying early scientific writing and provide implications for educational practices and curriculum design. The primary research question guiding this analysis is whether a hierarchical framework can explain how young children develop and use modal forms of representation in scientific writing.

Background literature

Generative learning theory as articulated by [Brod \(2021\)](#) and [Fiorella and Mayer \(2015\)](#) builds on the foundational work of [Osborne and Wittrock \(1985\)](#), who argued that sense-making is an internal process involving the active generation of relations between known and unknown information. This perspective positions language as a generative tool and learning as an active construction of meaning. Crucially, language extends beyond written text to include multiple modes of representation, each functioning as a semiotic resource that scaffolds meaning-making ([Kress, 2010](#)). The present study applies a cognitively derived hierarchical organizational framework (depicted in [Figure 1](#)) to examine how knowledge is structured within a tiered meaning-making system. This framework progresses from the broadest, most general conventions of communication to the most specific representational details, thereby guiding the cognitive processes involved in constructing and expressing scientific understanding across modalities ([Carroll, 1993](#); [Gobet et al., 2001](#); [Kintsch, 1998](#); [van Merriënboer and Sweller, 2010](#)). At the highest level, scientific discourse establishes the overarching norms



and structures of communication, influencing how learners engage with and make sense of science (Norris and Phillips, 2003). Within this discourse, genres define the communicative purpose and style, serving as scaffolds for structuring meaning through conventionalized multimodal configurations (Kress, 2010; Bateman, 2008; Bateman and Wildfeuer, 2014). In primary science education, informational genres predominate, supporting young learners in engaging with scientific explanations and phenomena (McNeill and Krajcik, 2012). Text structures such as description, sequence, and cause/effect provide rhetorical patterns that organize content and enhance coherence. Modes, including linguistic, visual, and graphical resources, then externalize and clarify knowledge, functioning as epistemic tools that support conceptual understanding (Ainsworth, 2014; Tang, 2020).

Composing multimodal texts, therefore, serves not only expressive purposes but also epistemic ones, providing a means through which learners generate, refine, and deepen understanding (diSessa, 2004). Young children's scientific text production is conceptualized as a cognitively demanding activity involving metacognitive planning (Flavell, 1979; Veenman et al., 2006), selective integration of information (Mayer, 2020; Brod, 2021), and the strategic use of mode affordances to externalize and reorganize knowledge (Ainsworth, 2006; Kress, 2010). Learners must choose and organize content, attend to coherence across semiotic resources, and use multimodal strategies to clarify or extend meaning. These practices reflect the central tenets of generative learning wherein learners select, organize, and integrate

information to produce lasting understanding (Brod, 2021; Fiorella and Mayer, 2015; Mayer, 2020), illustrating how multimodal composition functions as both a cognitive and epistemic activity. Below, key literature regarding how these elements contribute to early learners' development of representational skills is reviewed.

Discourse: the foundation of scientific communication

Scientific discourse refers to the complex system of conventions, symbols, and semiotic resources that enable knowledge construction and communication (Kress and van Leeuwen, 2021; Bateman, 2008). These resources include linguistic expressions, visual images, mathematical symbols, and data representations (van Leeuwen, 2005; Stöckl, 2004, 2020). Scientific discourse is characterized by its precision, structure, and adherence to established communicative norms that govern how scientific arguments are formulated, how evidence is presented, and how findings are validated (Kress, 2010). Critically, science education involves both learning scientific discourse and learning through scientific discourse (Lemke, 1998). Developing scientific discourse competency requires that learners understand the relationship between different semiotic resources and their role in scientific meaning-making (Norris and Phillips, 2003). Within this broad discourse, specific genres structure communication and define how scientific information is organized.

Genres in scientific discourse and the role of informational text

Genres in scientific discourse serve as scaffolds that direct the interpretation and production of scientific texts, shaping how information is organized, presented, and understood (Kress, 2010; Bateman, 2008). Bateman et al. (2017) further elaborate how genre operates within a layered architecture of multimodal documents, reinforcing the importance of analyzing how structural, visual, and linguistic features coalesce to shape meaning. These genres, which include research articles, review papers, lab reports, policy briefs, and popular science articles, adhere to distinct structural and rhetorical conventions (Hyland, 2015; Swales, 1990). The focus of this study is constrained to the genre of informational text, a key genre in early science education. Informational text is distinguished by its objective to inform and educate, often using a straightforward, expository style accessible to a wide audience, including young learners (Mayer, 2020; Duke, 2000).

Engaging with informational texts introduces children to scientific discourse conventions and supports the development of representational competencies (Palincsar et al., 2001; Fang, 2006). Informational texts also require learners to navigate multimodal representations, integrating textual, visual, and graphical information to construct meaning (Glushko, 2013). Research suggests that early and consistent exposure to informational texts strengthens comprehension and prepares learners for more advanced scientific literacy (Wexler, 2020). While extensive research has examined how children comprehend informational texts, less attention has been given to how they learn to produce them (Duke, 2000; van den Broek, 2010; Boucheix et al., 2020). This study addresses this gap by investigating how young learners develop the ability to produce informational scientific texts, focusing on how they assimilate and apply conventions of genre, text structure, and multimodal representations.

Text structure: organizing content within genre

Text structures in scientific discourse act as global frameworks that organize and provide coherence to informational text. These organizational patterns scaffold the integration of content and guide the selection and coordination of modes, influencing how ideas are communicated and interpreted. In expository science texts, the use of clear and recognizable structures enables learners to understand both what is being communicated as well as how to navigate relations among concepts. The ability to recognize and appropriately apply text structures is therefore fundamental to both comprehension and production in scientific writing (Meyer, 1975; Williams, 2018).

Informational texts can be categorized in different ways depending on the theoretical framework. In this study, we adopt a typology common in U.S.-based cognitive and developmental literacy research, which classifies informational texts into four primary types based on their communicative purposes: argumentative/persuasive, procedural, narrative non-fiction,

and expository (Duke and Tower, 2004; Purcell-Gates et al., 2007; NAEP, 2022). Of these, the expository type is particularly relevant to early science education, as it emphasizes the structured presentation of factual information about the natural world. This genre contrasts with the more socially grounded genre categorizations found in systemic functional linguistics (e.g., report, explanation, exposition) that focus on staged, goal-oriented social processes (Christie and Derewianka, 2008; Martin and Rose, 2012). Our approach aligns with U.S. literacy standards and instructional practices, including those used in the participating schools. In this study, students were asked to compose expository texts that would teach scientific concepts to a peer one grade level below. These texts typically adopt one of five common organizational patterns: definition/description, sequence, compare/contrast, cause/effect, and problem/solution (NAEP, 2022). Each of these structures serves as a scaffold for learners' attempts to communicate content clearly, often drawing on distinct constellations of linguistic and visual-graphical modes to support their communicative goals (Tippett, 2010; Halliday and Martin, 1993).

From a generative learning perspective, text structures serve as both cognitive and epistemic tools that help students organize new information, activate relevant prior knowledge, and construct coherent explanations (Wittrock, 1992; Fiorella and Mayer, 2015; Prain and Hand, 2016). Choosing an organizational format aligned with a communication goal (e.g., using a sequence to explain a process or a compare-and-contrast structure to analyze differences) supports learners in making conceptual connections and coordinating across multiple modes. Proficiency in using such structures improves students' ability to retain and convey scientific information by facilitating integration and elaboration, both of which are essential for knowledge construction (Bartlett, 1978; Meyer et al., 1980; Goldman and Rakestraw, 2000; Duke et al., 2021).

Importantly, the effectiveness of a given text structure often depends on how well it works in tandem with other modes of expression. Multimodal texts that align structural patterns with appropriate visual-graphical representations (e.g., diagrams that follow a sequence or charts that illustrate comparisons) are more likely to enhance comprehension and memory (Mayer, 2020). This interplay between text structure and mode selection is vital to the development of multimodal competency. As learners' metacognitive awareness and epistemic engagement increase, they become more strategic in coordinating structure and mode to communicate their scientific ideas with clarity and precision (Tang et al., 2014). In this way, the development of multimodal scientific writing is more than an ability to follow format.

Modes: the means used to convey scientific information

In the hierarchical organization of scientific communication, particularly within written texts, modes function as semiotic resources that link the abstract concept of discourse with the structured frameworks of genre and informational text. Defined by Jewitt et al. (2016) as "a set of socially and culturally

shaped resources for making meaning” (p. 9), modes are differentiated by their distinct affordances that signify both social and formal aspects of communication. The social aspect represents evolving representational needs within a community while the formal aspect relates to the functions a mode fulfills in communication.

Scientific discourse, particularly in written informational texts, most commonly draws upon verbal-linguistic (e.g., spoken and/or written text), visual-graphical (e.g., images, diagrams, graphs), and mathematical-symbolic (e.g., number systems, equations; Van Rooy and Chan, 2017). These three mode types are foundational in supporting scientific explanation and argumentation, and are key vehicles for communicating complex concepts in ways aligned with the expectations of the discipline. Other mode types including gestural-kinesthetic modes (e.g., hand gestures, gaze), and material-operational modes (e.g., physical objects, tactile manipulatives) play a significant roles in embodied science learning particularly in hands-on or spoken classroom contexts. Because these modes are not typically found in written form, they fall outside the scope of this study. Rather, the emphasis here is on how young learners use written linguistic modes and visual-graphical modes in their texts to convey scientific ideas, demonstrating their evolving multimodal competency and capacity to engage with scientific discourse (Lemke, 1998; Unsworth, 2001).

From a generative learning perspective, modes function not only as representational resources but also as epistemic tools; that is, they support learners in actively constructing, refining, and externalizing their understanding (Fiorella and Mayer, 2015; Prain and Hand, 2016). This dual role highlights how modes facilitate both communication and cognition: when learners select and coordinate modes to represent their knowledge, they are engaging in a process of meaning-making that involves selecting relevant information, organizing it coherently, and integrating it with prior knowledge. In this way, modes do not simply transmit information; rather, they shape how ideas are conceptualized in the first place, and serve as active components of knowledge building.

Linguistic modes

Linguistic modes effectively structure content, emphasize critical details, and support both comprehension and recall (Fisher et al., 2012; Graham et al., 2020). Core linguistic features such as titles, headings, and lists outline the overarching structure of a text, segment information into digestible units, and cue readers about transitions between ideas. These elements reduce cognitive load and promote coherence, making them especially important for novice writers and readers (Hartley, 2014; Pettersson, 2019). Supplementary linguistic features like tables of contents, indexes, and glossaries enhance navigability and concept recall, especially in more extended or structured scientific texts (Harp and Mayer, 1998; Sweller, 2011). Within a generative learning framework, these linguistic tools do more than guide the reader; they also help the writer make sense of what they are trying to communicate. For young learners, constructing titles or headings requires abstraction,

synthesis, and prioritization of information, generative skills that support deeper understanding and concept reorganization.

Visual-graphical modes

Visual-graphical modes including illustrations, diagrams, tables, and maps, complement, enhance, or extend written information (Carney and Levin, 2002; Fingeret, 2012). These modes are particularly powerful in science because they support visual perception, direct attention to key features, and promote spatial and relational reasoning (Boucheix et al., 2020; Fingeret, 2012; Pettersson, 2019). The choice of visual-graphical modes varies depending on the information being represented. Photographs ground abstract concepts in observable phenomena, while diagrams simplify systems or processes. Tables and graphs are typically used to present numerical data, and cycle diagrams or flow maps are appropriate for showing sequences or processes (Slutsky, 2014; Franconeri et al., 2021). Visual elements are often supported by labels and captions that help clarify meaning and ensure information is appropriately interpreted (Lamberski and Dwyer, 1983; Beck, 1984; Mayer, 2020).

Visual-graphical modes also operate as epistemic tools, especially when learners generate the visuals themselves. In drawing or designing a diagram, students must determine which features are important, how to abstract and simplify information, and how to represent relationships between components. These choices activate essential generative processes including organizing and elaborating on content that reinforce learning through the construction of external representations.

Synergistic integration of modes and content

The synergistic integration of linguistic and visual-graphical modes is critical for effective scientific communication and represents an important milestone in the development of multimodal competency. When used together, these modes enable learners to externalize internal representations in more complete and communicatively effective ways. Well-integrated multimodal texts are more than simple composites of text and image. They reflect deliberate decisions about what information to present in each mode and how the modes work together to support understanding. This process mirrors the organizing, transforming, and connecting functions central to generative learning (Wittrock, 1992; Mayer, 2020). As learners develop metacognitive awareness of how different modes function including how they can be used in concert, they begin to make increasingly strategic decisions about representation. This progression enhances their ability to communicate ideas clearly while concomitantly deepening their understanding of the content itself. In this way, modes become tools for constructing and refining scientific knowledge.

Content: the core of scientific knowledge

Content learning in science, particularly when producing scientific texts, embodies an active and constructive process. Learners do not merely receive information they also actively reshape and transform it to build a meaningful understanding

consistent with the principles of generative learning theory. Learning is optimized when students are involved in selecting relevant content, organizing it meaningfully, and integrating it with existing knowledge structures (Wittrock, 1974; Fiorella and Mayer, 2015; Hand et al., 2021). When young learners write about science, they are required to analyze and synthesize new ideas, distill them into key points, and express them in their own words. These acts of summarizing, explaining, comparing, and categorizing are both communicative and deeply generative. They prompt students to construct internal mental models that reflect both their prior experiences and the cognitive demands of the task (Halford, 2014). In turn, these mental representations serve as scaffolds for future learning, supporting the formation of new connections and strengthening their conceptual frameworks (Pilegard and Fiorella, 2016; Wylie and Chi, 2014).

Importantly, when learners express content through writing and other modes, these expressions function as epistemic tools that make thinking visible and subject to reflection, revision, and elaboration (Prain and Hand, 2016). In this way, content production, especially when guided by generative strategies, goes beyond the simple transmission of ideas and becomes a mechanism for conceptual transformation. This claim was supported by empirical evidence from Hand et al. (2021), who reviewed 81 intervention studies across K-16 settings and found consistent improvements in science learning when students were engaged in writing-to-learn activities that required them to actively structure and represent content. These findings emphasize how content, when transformed through strategic representation, can drive knowledge construction by enabling learners to externalize, refine, and extend their understanding of scientific ideas.

Multimodal competency: integrating discourse, genre, text structure, mode, and content

Multimodal competency, particularly in the context of primary school science education, equips learners with the ability to construct and interpret meaning across diverse representational forms. This aligns with Bateman and Wildfeuer's (2014) emphasis on the systematic coordination of semiotic modes within complex documents, where layout, genre, and rhetorical structures co-construct meaning across representational systems. It also reflects perspectives from systemic functional linguistics, which conceptualize language and multimodal texts as social semiotic systems through which knowledge is constructed and communicated (Halliday and Martin, 1993; Schleppegrell, 2004). From this view, scientific texts are more than neutral vessels of information; rather, they are structured forms of discourse shaped by discipline-specific purposes, genres, and representational demands.

Multimodal competency also supports a richer, more integrated grasp of scientific literacy by enabling students to coordinate elements of discourse, genre, text structure, mode and content in ways that enhance both comprehension and communication (Ainsworth, 2006; Lemke, 2004). These skills are developed through individual cognitive processing as well as through participation in socially situated literacy practices (Christie

and Derewianka, 2008; Humphrey, 2018; Chen et al., 2020). As students engage in composing scientific texts, they draw on shared cultural tools including diagrams, labeled illustrations, captions, and comparative text structures that reflect the epistemic norms of the discipline (Unsworth, 2001; Bezemer and Kress, 2008). Developing this competency allows young learners not only to express what they know but also to refine and extend their thinking, as they make deliberate choices about how to represent complex scientific concepts across modes.

Embedding multimodal learning strategies into early science education strengthens conceptual understanding and communication skills by promoting engagement, supporting meaning-making, and aligning with varied cognitive strengths among learners (Tang et al., 2014). When students are guided in how to integrate linguistic and visual-graphical modes with coherent structures and relevant content, they are more likely to engage in generative learning including selecting, organizing, and elaborating on information in ways that deepen their scientific understanding (Fiorella and Mayer, 2015). These representational strategies also function as epistemic tools, enabling learners to externalize and evaluate their thinking and make connections between internal knowledge structures and external scientific discourse (Prain and Hand, 2016; Tang, 2020). Despite its recognized importance, the development of multimodal competency in early childhood education remains underexplored, particularly with regard to how it unfolds across multiple components of scientific text production. Much of the existing research has focused on isolated elements of multimodality like the role of visuals in comprehension or the function of language in constructing explanations without fully examining how these elements interact within a broader representational system (Prain and Tytler, 2022; Waldrip et al., 2010).

In this study, we define hierarchical organization not merely as increasing complexity, but as a structured coordination of meaning-making elements, where higher-order conventions (e.g., discourse and genre) constrain and shape lower-level features (e.g., text structure, mode, and content). This perspective draws from multimodal discourse theory (Kress and van Leeuwen, 2021; Bateman, 2008) and cognitive load theory (Sweller, 2011), suggesting that well-organized texts reflect deliberate decisions about how to present content for maximum coherence and communicative effectiveness. Hierarchical organization, then, reflects not just the presence of multiple modes, but the integration of those modes into a layered structure that supports both comprehension and expression. This framework guided our identification of distinct multimodal typologies that represent progressively more organized and integrated forms of scientific writing. These include Unlabeled Visual Texts, Labeled Visual Texts, Foundational Multimodal Texts, Text-Centric Narrative Texts, Redundant Complex Multimodal Texts, and Complementary Complex Multimodal Texts, which we present in order from least to most hierarchically organized.

The present study

The present study addresses this gap by applying a latent class analysis approach to examine the typologies of multimodal

competency found in young learners’ scientific texts. Rather than focusing on modal features in isolation, we analyze the interaction of text structure, linguistic and visual-graphical modes, and scientific content within a hierarchical organizational framework. This approach enables us to delineate the varied strategies primary school children use to understand, organize, and communicate scientific ideas and to identify the underlying cognitive processes including metacognitive awareness, selective integration, and sensitivity to mode affordances that support these strategies. In doing so, this study contributes to a more comprehensive understanding of how multimodal competency develops during the early years of science learning and how it supports both the communicative and epistemic functions of writing.

Methods

Research design and sample

A quantitative content analysis methodology was used to examine primary school children’s production of science-based summary writing samples. This study employed a statistical approach, including latent class analysis (LCA), to derive data-driven typologies based on coded attributes. Content analysis is defined as a systematic, objective method for quantifying and describing phenomena within written, spoken, or visual communication forms, allowing for the numerical evaluation of specific attributes or characteristics in the students’ writing samples, such as frequency, presence, and relations among text structure, mode, and content elements (Krippendorff, 2019; Skalski et al., 2019). Because the goal of this study was to investigate how students use writing as a representational and meaning-making practice, this approach aligns with prior work emphasizing scientific texts as epistemic tools and written representations as externalizations of internal cognitive processes (Fiorella and Mayer, 2015; Prain and Hand, 2016).

Sample source and composition

Across the 4-year timeframe (2014–2017), a total of 1,705 summary writing samples were collected from 1,008 students, resulting in an average of 1.69 samples per student ($SD = 0.81$). Samples were distributed across kindergarten ($n = 592$), first grade ($n = 579$), and second grade ($n = 534$). Each sample was categorized by its scientific content focus: plants, animals, or matter. As shown in Table 1, the majority of kindergarteners wrote about plants, first graders about animals, and second graders about matter, in alignment with state curricular expectations.

Procedure

As part of a broader investigation into the efficacy of various science instructional methods, teachers were tasked with collecting writing samples each March following a professional development session. This session emphasized the importance of multimodal representation and scientific writing as both a communicative practice and as a cognitive and epistemic process through which learners could construct scientific understanding. Teachers had 2

TABLE 1 Distribution of writing samples by grade level and scientific topic.

Grade level	Plants	Animals	Matter	Total
Kindergarten	591	1	0	592
1st Grade	38	414	127	579
2nd Grade	1	139	394	534
Total	630	554	521	1,705

Topic distribution reflects grade-level alignment with science curricula used in participating schools during the data collection period (2014–2017).

weeks to teach a lesson and collect a sample from each student. Samples were produced in a single session after students completed their most recent science unit and were digitized for coding and analysis.

Teacher professional development for summary writing

Participating teachers engaged in a 4-day professional development program focused on implementing the Science Writing Heuristic (SWH) approach (Hand, 2008). This school-initiated initiative involved all K–2 teachers across participating schools and occurred during the summer prior to data collection. As part of this program, teachers participated in a researcher-led 30-min training session (without compensation) that emphasized the use of multimodal scientific writing to support both comprehension and explanation in early science learning.

The training extended teachers’ existing knowledge of language-focused instruction by integrating a focus on science-specific language practices and representational tools. Teachers were introduced to the idea that multimodal texts (i.e., text combining written, visual, and graphic modes) can serve not only as vehicles for communication but also as cognitive tools for deepening scientific understanding. *The session highlighted how scientific texts use both language and visuals to explain phenomena and convey ideas, with specific attention to genre (informational/expository), common organizational structures (e.g., compare/contrast, description), and the affordances of diagrams, labels, and illustrations.* To model these principles, teachers were shown a demonstration lesson comparing unimodal texts (text-only) with multimodal texts that incorporated diagrams, labels, and organizational features. This example illustrated how multimodal texts can better support comprehension and communication of complex scientific content. Teachers were then asked to design a lesson in which students created their own scientific texts for a peer one grade level below, using multiple modes to make ideas more accessible. Materials used during the session are available from the last author.

All participating teachers were responsible for teaching science as part of their state-mandated curriculum. In accordance with state and district policies, K–2 students received instruction in physical science, life science, and earth and space science over the course of the school year. However, the total time allocated for science instruction was typically limited to 30–45 minutes per week. While explicit instruction in multimodal scientific representation varied

across classrooms, the training described above served as a common anchor point. Students were introduced to the idea that scientific ideas can be represented in multiple ways and were encouraged to use features such as labeled diagrams, comparisons, and organizational text structures in their writing. Teachers incorporated these ideas into their regular instructional routines based on their own teaching contexts and curricular priorities.

Although we did not systematically observe classroom instruction in this study, the combination of professional development, curricular standards, and the sample prompts used suggest that students had some exposure to genre-specific features of scientific texts and to the value of combining modes to enhance communication. Future work will benefit from more detailed classroom observation and documentation of instructional practices.

Coding schemes

Three separate coding schemes were developed and applied to the full dataset. Each scheme was supported by training procedures and interrater reliability checks. All codes were defined in advance of LCA and applied systematically to the full dataset using structured protocols. The first focused on identification of the informational text structures that students used to frame their samples; the second focused on the linguistic and visual-graphical modes students embedded in their samples; and the third focused on the specific content that students presented in their samples. Table 2 includes all codes, the prevalence of each code across all samples, and a definition for each code.

Text structures

Description. This coding scheme identified the informational text structures used by students. Because structure reflects how students select, organize, and sequence information, it serves as a window into the generative processes involved in scientific meaning-making (Wittrock, 1992). Text structures scaffold students' written expression, allowing them to express relationships among ideas in ways aligned with scientific communication practices.

Reliability. Samples were coded individually using the text structure definitions presented in Table 2. To establish reliability, a subset of 208 samples across the three topic areas was independently double-coded. The overall percent agreement was 98.1%, with a Cohen's kappa value of 0.97, indicating a high level of agreement (Landis and Koch, 1977). For plants, percent agreement across all three text structures was 95.8% (Cohen's kappa = 0.92); for animals, percent agreement was 97.6% (Cohen's kappa = 0.96); for matter, percent agreement was 100% (Cohen's kappa = 1.00).

Data Organization. Descriptive analyses indicated predominant use of description (68.3%), compare/contrast (28.0%), and sequence (3.6%) structures, with notable variations across grade levels. No cause/effect or problem/solution structures were identified.

Linguistic and visual-graphical modes

Description. This coding scheme examined how students employed linguistic and visual-graphical modes in their scientific explanations. These modes serve as the material enactment of epistemic tools, externalizing internal understandings and enabling effective communication. The distinction between primary and auxiliary linguistic and visual-graphical features captures the strategic integration of modes consistent with multimodal generative learning (Ainsworth, 2006; Fiorella and Mayer, 2015).

Reliability. Two graduate students in science education, in collaboration with the lead author, designed the coding scheme for these linguistic and visual-graphical modes. They also developed training videos and materials. Three undergraduate students were then trained using these resources to complete the coding task. Initial consensus between the lead author and the students was reached on 20 samples, with any disagreements resolved through discussion until a consensus was reached. These coders then independently assessed the remaining 1,685 samples. For reliability, a random subset of 161 samples was double-coded, yielding a 94.3% agreement rate across coders (Cohen's kappa = 0.94).

Data Organization. Data were aggregated into primary and auxiliary linguistic modes and visual-graphical modes based on features commonly found in early science texts. Primary linguistic features included elements such as titles, headings, prefaces, conclusions, and emphasized text, which helped structure and highlight key information. Auxiliary linguistic features captured navigational aids such as glossaries, sidebars, and tables of contents, though these were relatively rare. Visual-graphical features included labeled and unlabeled illustrations, diagrams, tables, and charts, along with supporting devices such as labels, captions, and connectors that linked visuals to accompanying text. These features were chosen for their cognitive and communicative affordances in science texts (Fingeret, 2012; Pettersson, 2019). Descriptive statistics for each coded mode are provided in Table 3. For the purposes of latent class analysis, each mode was recoded into a binary variable indicating its presence or absence in a given sample.

Scientific content

Definition. This coding scheme assessed the disciplinary content conveyed in students' writing and visuals. Because generative learning requires learners to reformulate and apply disciplinary knowledge, the scientific ideas expressed via text, visuals, or both provide insight into how students are actively constructing and communicating understanding (Wylie and Chi, 2014). Rubrics were created for each of three focal science topics: plants, animals (both living animals and fossils), and matter, based on the NGSS (NGSS Lead States, 2013). Coding rubrics for each topic are included in the Appendix. Corresponding coding manuals provide an overview of relevant standards, visual exemplars from the sample pool, and coding rules. These manuals are available upon request from the lead author.

Reliability. Life Sciences content fit into three broad standards: (1) Structure and Processes: From Molecules to Organisms; (2) Heredity: Inheritance and Variation of Traits; (3) Ecosystems: Interactions, Energy and Dynamics. Physical Sciences content

TABLE 2 Coding schemes, codes, percentage of all samples, and definitions.

Coding scheme Type codes	% All	Definition
Text structures		
Description or definition	68.3%	Major ideas are presented by outlining characteristics, features, attributes, or examples.
Sequence, chronological, process	3.63%	Major ideas or events are arranged in the order in which they occur.
Compare/contrast	28.0%	Details, descriptions, or examples explain how two or more major ideas are similar or different.
Cause/effect	–	Major ideas are presented as causes or effects, supported by details explaining their origins or outcomes.
Problem/solution	–	Major ideas are framed as problems, accompanied by details elucidating potential solutions.
Primary linguistic modes		
Titles	40.6%	A short summary of the content of the entire text.
Headings subheadings	15.2%	A heading is a short summary of the content of a portion of the text (e.g., paragraph, chapter) while a subheading falls below a heading and divides the section into more specific categories.
Lists	9.7%	Bulleted/numbered text lists that reduce long or complicated sentences, emphasize key points and concepts.
Emphasized text	19.5%	Stylistic variations in text that signal a word or phrase is important (e.g., bold, italics, color, font size, all CAPS, underlining). All emphasized text in primary text (16.5%), auxiliary text (1.9%), and auxiliary graphic features (4.9%) are combined in this category.
Questions	5.7%	Linguistic devices in the form of a question used to assist and stimulate learning or encourage thinking about presented material. All questions in primary text (4.8%), auxiliary text (1.1%), and auxiliary graphic features (0.1%) are combined in this category.
Preface	13.0%	Statement that sets a purpose for reading or provides an overview of the content.
Conclusion	1.1%	A finishing statement that indicates it is the end of the text or that summarizes/restates the main purpose.
Auxiliary linguistic modes		
Table of contents	0.5%	A list that provides the major chapters or sections in a text along with location (e.g., page numbers).
Index	0.1%	Typically at the end of a text that indicates topics and locations in the text (e.g., page numbers).
Glossary	5.0%	An alphabetical list of key terms with definitions found at the end of a text, a chapter, or section of text.
Pronunciation guide	0.5%	Indicate how to pronounce key words.
Sidebar	5.2%	Boxes of information often placed near the side of the main text or column(s).
Appendix	0%	A section at the end of text that provides additional information that is tangential to the main text.
Footnotes references	0%	May be explanatory or sources/citations.
Visual-graphical modes		
<i>Objects or people</i>	94.0%	Visual representations that are redundant with text content
Photographs	0%	An exact image of what a person, animal, place, or thing looks like
Unlabeled illustrations	34.4%	A simple drawing of what a person, animal, place, or thing looks like, does not extend information contained in the text; with or without labels
Labeled illustrations	30.6%	
Unlabeled diagrams, all	9.8%	A graphic that identifies the parts of a subject and their spatial or functional relationship; emphasizes details; or shows dimensions. With or without labels

(Continued)

TABLE 2 (Continued)

Coding scheme Type codes	% All	Definition
Labeled diagrams, all	37.5%	
Rebus	2.0%	A visual icon placed above the written text it depicts
Breakdown of diagram types		
Part labels	31.1%	A diagram where labels identify individual parts of the visual
Whole labels	4.3%	A diagram where labels identify whole parts of the visual
Magnification	4.3%	A close-up view of a person, animal, place, or thing
Cross-sections	0.2%	A graphic that presents an inside view of an animate or inanimate object
Comparisons	0.8%	A graphic that depicts a person, animal, place, or thing by comparing it to something familiar
Overlays	0%	A graphic placed over another graphic to add information or show a relation between two things
Numbers, quantities, and categories	15.8%	Visual representations of data used to organize information to highlight patterns and relations
Tables	14.1%	A list of facts or figures systematically displayed in rows and columns
Charts/Graphs	2.3%	A graphic that is a visual representation of data showing how two or more sets of qualities or numbers are related to each other
Order	5.4%	Chronologies, steps, or processes
Timeline	0%	A graphic that depicts a chronological arrangement of events
Life Cycle	5.4%	A graphic that depicts the different stages in the life of an organism
Spatial content	0.4%	Reference (i.e., location of geographic boundaries, physical features, cultural features) or thematic representations (i.e., variation of topics across a geographic area)
Maps	0.4%	A graphic that displays geographical, sociological, or scientific information using a spatial representation of an area
Visual-graphical modes—supplemental		
Labels	55.6%	Text that identifies specific elements or features on a graphic. Labels can be placed directly on the graphic, or around the graphic and connected with a line or arrow to the relevant part.
Captions	34.2%	Short phrases or sentences underneath or beside photos, illustrations, or other graphics that describe the graphic as a whole
Connectors	57.6%	Arrows or lines that link caption or label text to a graphic or that shows movement or a process
Table headings	2.6%	Text entries, presented as column or row titles, that support scanning the table to find correct or desired information
Thought bubbles	3.4%	A rounded or cloud-shaped outline containing words that represent a person or animal's thoughts.
Scientific content		
Written text only	76.5%	Content that exclusively uses written language to convey scientific information without the support of visual elements.
Visual only	83.9%	Content that solely relies on visual elements such as diagrams, charts, and illustrations to represent scientific concepts without accompanying textual descriptions.
Mixed written text and visual	51.7%	Content that integrates both written language and visual elements to present scientific information using a multimodal approach.

Definitions for text structures adapted from Goldman and Rakestraw (2000), Hess (2010), Bovair and Kieras (1991), Kobayashi (2002), and Meyer (1975). Definitions for linguistic and graphical-visual modes adapted from Fingeret (2012), Guo et al. (2018), Hess (2010), Hartley (2014), Pettersson (2019), Roberts et al. (2013), and Davis and Hunkins (1966)).

TABLE 3 Descriptives and chi-squares contrasting typologies by linguistic and visual-graphical modes.

Category code	Summary writing sample typologies							
	Complementary complex multimodal	Redundant complex multimodal	Foundational multimodal	Unlabelled visuals	Text-centric narrative	Labeled visuals	Significance test	Effect size
Linguistic modes—primary								
Titles	83.0% ^{a+}	68.1% ^{b+}	47.1% ^c	10.5% ^{d−}	57.7% ^{bc+}	5.6% ^{d−}	528.23***	0.56
Headings/subheadings	35.7% ^{a+}	41.6% ^{a+}	1.9% ^{bc−}	2.9% ^{c−}	19.8% ^d	0% ^{b−}	381.49***	0.47
Lists	18.8% ^{ab+}	18.2% ^{b+}	9.6% ^a	0.3% ^{c−}	17.8% ^{ab+}	0% ^{c−}	128.80***	0.28
Emphasized text	39.3% ^{ab+}	42.9% ^{b+}	26.3% ^{ac+}	0% ^{d−}	16.6% ^c	1.2% ^{d−}	317.86***	0.43
Questions	15.2% ^{a+}	15.5% ^{a+}	2.2% ^{bcd−}	0.3% ^{d−}	4.7% ^c	0.6% ^{bd−}	127.14***	0.27
Preface Conclusion	49.1% ^{a+}	31.6% ^{b+}	1.3% ^{c−}	0.6% ^{c−}	19.0% ^{d+}	1.2% ^{c−}	360.80***	0.46
Linguistic modes—auxiliary								
Navigation and comprehension tools	17.0% ^{a+}	15.5% ^{a+}	2.6% ^{bc−}	0% ^{cd−}	4.7% ^b	0% ^{d−}	139.74***	0.29
Sidebars	10.7% ^{a+}	11.3% ^{a+}	6.1% ^a	0% ^{b−}	6.3% ^a	0% ^{b−}	71.51***	0.21
Visual-graphical modes								
Illustrations, unlabelled	31.3% ^a	41.3% ^{a+}	14.4% ^{b−}	62.3% ^{c+}	60.1% ^{c+}	1.5% ^{d−}	410.00***	0.49
Illustrations, labeled	85.7% ^{a+}	74.0% ^{a+}	13.8% ^{b−}	3.2% ^{c−}	25.3% ^d	9.6% ^{b−}	717.03***	0.65
Diagrams, unlabelled	8.9% ^{ab}	9.1% ^b	2.9% ^{ac−}	32.6% ^{d+}	2.8% ^{ac−}	1.5% ^{c−}	242.22***	0.38
Diagrams, labeled	17.0% ^{a−}	13.9% ^{a−}	87.2% ^{b+}	1.0% ^{c−}	3.6% ^{c−}	83.0% ^{b+}	1042.60***	0.78
Rebus	2.7% ^{ab}	5.1% ^{b+}	0.6% ^a	1.0% ^a	2.4% ^{ab}	0.3% ^{a−}	28.50***	0.13
Tables	37.5% ^{a+}	39.9% ^{a+}	0.3% ^{b−}	0% ^{b−}	19.0% ^{c+}	0% ^{b−}	418.34***	0.50
Charts/graphs	8.9% ^{a+}	5.1% ^{a+}	0.3% ^{bcd−}	0% ^{d−}	3.6% ^{ac}	0% ^{bd−}	57.80***	0.18
Cycle diagrams	5.4% ^{ab}	9.4% ^{b+}	0.6% ^{c−}	4.2% ^{abc}	3.2% ^{ac}	8.2% ^{ab+}	34.08***	0.14
Maps	0% ^a	1.3% ^{a+}	0.6% ^a	0% ^a	0% ^a	0% ^a	12.50*	0.09
Visual-graphical modes—supplemental								
Labels	52.7% ^a	52.5% ^a	99.7% ^{b+}	1.3% ^{c−}	19.4% ^{d−}	96.2% ^{e+}	984.40***	0.76
Captions	92.0% ^{a+}	80.4% ^{a+}	23.7% ^{b−}	3.5% ^{c−}	25.7% ^{b−}	8.8% ^{c−}	772.97***	0.67
Connectors	83.9% ^{a+}	66.5% ^{b+}	98.7% ^{c+}	0.3% ^{d−}	22.9% ^{e−}	79.8% ^{a+}	874.05***	0.72
Table headings	9.8% ^{a+}	6.4% ^{a+}	0% ^{b−}	0% ^{b−}	4.0% ^a	0% ^{b−}	71.30***	0.20
Thought bubbles	2.7% ^{ab}	8.0% ^{b+}	3.8% ^{ab}	1.6% ^a	2.0% ^a	0.9% ^{a−}	36.12***	0.15

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; Cramer's $V = \leq 0.2$ is small; 0.2–0.6 = moderate; ≥ 0.6 = large; when superscript letters in a particular row differ (i.e., one column is a and the other column is b, the column proportions for those two comparisons differ significantly at the $p < 0.05$ level. + The standardized residual was > 1.96 indicating that the number of cases in this cell was significantly greater ($p < 0.05$) than expected; − The standardized residual was less than -1.96 indicating that the number of cases in this cell was significantly fewer ($p < 0.05$) than expected.

fit one standard: Matter and Its Interactions. Initial consensus scoring for 20 samples involved five undergraduate students and the lead author, with disagreements resolved through discussion. The team then independently coded the remaining 1,685 samples. Reliability checks on a random subset of 211 samples distributed approximately equally across topics showed an overall percent agreement of 93.7%, with a Cohen's kappa of 0.93, indicating good agreement (Landis and Koch, 1977). For plants, percent agreement was 94.6% (Cohen's kappa = 0.95); for animals, percent agreement was 93.8% (Cohen's kappa = 0.94); for matter, percent agreement was 94.5% (Cohen's kappa = 0.94).

Data Organization. Content data were categorized into three types: depicted in the written text only, in the visual content only, or in both the text and the visual content. For mixed content, all instances where the content was found in both written text and visuals were scored and then doubled to account for the integration of text and visuals (e.g., learners who drew a picture of a flower with a stem, leaves, and roots and then included text labels for the stem, leaves, and roots would get 6 mixed content units), thereby acknowledging the added complexity and effort in conveying the same content in both text and visual form. These three composite variables were summed to create a total content score. Then, the percentages for each content type relative to the total amount of content presented were calculated by dividing the amount of each content type by the total content and multiplying by 100. These details are provided in Table 4.

Analytical approach

A combination of descriptive statistics, chi-square tests, and mixed-effects models were used to examine how students integrated structure, mode, and content. These strategies captured variation in multimodal scientific communication across topic, grade, and macrostructure. Some modes were found infrequently in the samples, rendering it impossible to include each as individual variables. When this occurred (i.e., comprised fewer than 2% of all samples), they were grouped with similar modes. Separately, emphasized text and questions were aggregated across primary text, auxiliary text, and graphics. Prefaces and conclusions were combined. All navigation and comprehension tools were combined into one variable (i.e., table of contents, index, glossary, pronunciation guide). All diagrams except cycle diagrams were combined into two variables depending on the presence or absence of labels. Illustrations were separated by the presence or absence of labels. Tables, charts, and graphs were combined into one variable. Cycle diagrams and maps were combined into one variable.

Multimodal competency typologies

Multimodal competency typologies were created by looking for patterns or sub-groups associated with students' integration of text structure, linguistic and visual-graphical modes, and content. A Latent Class Analysis (LCA) was used to identify these typologies or latent classes (LCA; Muthén and Muthén, 2000). LCA is an analytical approach that categorizes individuals

based on observed variables, explaining them through unobserved or latent categories. The LCA included binary indicators that identified the presence or absence of primary and auxiliary linguistic modes and primary and auxiliary visual-graphical modes as well as the percentage of content units depicted in different modes (i.e., written text content only, visual content only, and both written text and visual content). Based on this analysis, samples were assigned to specific latent classes according to their likelihood of belonging to each class, determined by their posterior class membership.

In all latent class models, 100 random starts were initiated to avoid local solutions. To evaluate the statistical fit of each k-class solution, various approximate fit indices and likelihood ratio tests were used. The fit indices included Bayesian Information Criterion (BIC), Sample-Size Adjusted Bayesian Information Criterion (SABIC), Akaike Information Criterion (AIC), and Approximate Weight of Evidence criterion (AWE). Lower values in these indices signify a better fit compared to a k-1 class model. Bayes Factor (BF) and correct model probability (cmP) were also calculated for further model validation. The adjusted Lo-Mendell-Rubin (LMR) and bootstrap likelihood ratio test (BLRT) were also examined such that significant *p*-values in these tests indicated a model fit superior to the k-1 class solution. The entropy index served as a measure of the clarity with which students were classified into various classes. Entropy values exceeding 0.80 are considered indicative of "good" classification (Nylund-Gibson and Choi, 2018). The Average Posterior Probability (AvePP) was also used as a measure of how well the model classified observations into their most likely class. Here, values >0.70 signal well-defined classes (Nagin, 2014). Finally, the model solutions were assessed for their theoretical and conceptual relevance, along with a qualitative examination of the conditional response probabilities of each latent class for all specified characteristics in the LCA (Masyn, 2013). The Appendix includes three supplemental figures (SF) with examples of writing samples by text structure, topic, grade, and multimodal competency level.

SPSS version 29 was used for data management and computation of descriptives. Stata/SE version 18.0 was used to perform mixed-effect regressions. M-Plus version 8.8 was used to compute the LCA. This study's design and its analyses were not pre-registered.

Results

Multimodal competency typologies

In the selection of the optimal model solution for the LCA, a comprehensive evaluation of fit indices and statistical tests was completed to ensure the chosen model accurately reflected the underlying structure of the data (Table 5). A 6-class typology model was selected over alternative solutions based on the following evaluation. The preference for lower values in BIC, SABIC, AIC, and AWE indices, which indicate a better fit to the data, supported the selection of the 6-class model. While the VLMR-LRT *p*-value suggested a less clear distinction for adding an additional class beyond the fifth, the significant improvement

TABLE 4 Descriptives contrasting typology by content, text structure, grade, and unique modes.

Category group		Summary writing sample typologies							
		Complementary complex multimodal	Redundant complex multimodal	Foundational multimodal	Unlabeled visuals	Text-centric narrative	Labeled visuals	Significance test	Effect size
Content ¹	All content (total/SE)	10.56 (0.75) ^c	27.58 (1.03) ^b	17.66 (0.71) ^c	10.79 (0.45) ^a	4.99 (0.26) ^d	17.92 (0.69) ^e	865.31***	12,901.52 ^p 12,219.17
Written text only (%)		23.4% ^a	16.9% ^{bd}	20.2% ^a	14.6% ^{cd}	96.1% ^e	7.9% ^f	1,411.47***	0.80 ^η
Visual only (%)		75.3% ^a	24.2% ^b	47.0% ^g	50.3% ^{dg}	3.2% ^e	66.0% ^f	302.43***	0.47 ^η
Combined text and visual (%)		1.4% ^a	55.5% ^b	31.3% ^c	24.3% ^{dg}	0.8% ^e	25.9% ^{fg}	223.15***	0.40 ^η
Text structure	Description (%)	23.2% ^{a−}	32.2% ^{a−}	89.1% ^{b+}	89.1% ^{b+}	58.5% ^c	91.8% ^{b+}	662.15***	0.44 ^ν
Sequence (%)		1.8% ^{ab}	3.5% ^{ab}	1.0% ^{b−}	4.8% ^{ab}	1.2% ^{b−}	7.6% ^{a+}		
Compare/contrast (%)		75.0% ^{a+}	64.3% ^{a+}	9.9% ^{b−}	6.1% ^{b−}	40.3% ^{c+}	0.6% ^{d−}		
Grade	Kindergarten (%)	2.7% ^{ab−}	0% ^{c−}	0.6% ^{bc−}	76.0% ^{d+}	5.1% ^{a−}	98.2% ^{e+}	2,055.69***	0.78 ^ν
1st Grade (%)		0% ^{a−}	24.7% ^{b−}	93.9% ^{c+}	21.4% ^{b−}	48.6% ^{d+}	1.2% ^{a−}		
2nd Grade (%)		97.3% ^{a+}	75.3% ^{b+}	5.4% ^{c−}	2.6% ^{cd−}	46.2% ^{e+}	0.6% ^{d−}		
Total linguistic modes		5.96 (0.62) ^a	6.28 (0.36) ^a	1.89 (0.13) ^b	0.23 (0.03) ^c	3.25 (0.24) ^d	0.12 (0.02) ^e	1,039.80***	6,898.89 ^p 5,807.85
Total visual-graphical modes		16.02 (0.98) ^a	13.31 (0.46) ^b	12.44 (0.47) ^c	1.45 (0.09) ^d	4.50 (0.22) ^e	9.77 (0.36) ^f	1,490.44***	11,048.34 ^p 9,834.08

*** $p < 0.001$ Model | Significance Test | Effect Size: ^p = mixed effects negative binomial regression | Wald χ^2 | AIC null model, AIC full model; ^η = ANOVA | F-test | partial η^2 ; ^ν = Chi-Square | χ^2 | Cramer's V; Note. When superscript letters in a particular row differ (i.e., one column is a and the other column is b, the column proportions for those two comparisons differ significantly at the $p < 0.05$ level. ⁺ The standardized residual was >1.96 indicating that the number of cases in this cell was significantly greater ($p < 0.05$) than expected; [−] The standardized residual was less than -1.96 indicating that the number of cases in this cell was significantly fewer ($p < 0.05$) than expected. ¹Content percentages do not total to 100% as some samples contained no content.

in other indices for the 6-class model, coupled with a high entropy value, supported its superior classification accuracy and model fit.

The model’s robustness in classifying observations into well-defined classes was sound. Benchmarks indicate that the entropy index should exceed 0.80 (it was 0.94), and the AvePP values should be equal to or >0.70 (all exceeded 0.93) for the 6-class solution. These metrics, indicative of “good” classification quality, are consistent with recommendations by Nylund-Gibson and Choi (2018) and Nagin (2014), emphasizing the importance of clear and distinct class separations in LCA. Moreover, the theoretical and conceptual relevance of the 6-class model was carefully considered, ensuring that the identified classes were statistically distinct, meaningful, and interpretable within the context of primary grade children’s multimodal competency.

Evidence supports the use of these comprehensive criteria for model selection in LCA to balance statistical fit with theoretical interpretability to ensure that the chosen model accurately captures the complexity of the data while remaining grounded in the substantive research question (Masyn, 2013). The decision to adopt the 6-class model, despite the mixed signal from the VLMR-LRT *p*-value, was justified by the overall improvement in fit indices, the high entropy value, and the model’s alignment with the conceptual framework of the study.

To facilitate interpretation of the six emergent typologies, we organized them according to the degree of hierarchical organization observed across text structure, mode use, and content integration. This ordering, from least to most integrated, aligns with the theoretical framework described earlier, in which the coordination of discourse, genre, structure, modes, and content reflects increasing multimodal sophistication. Typologies that demonstrated minimal integration of modes and content, such as Unlabeled Visual Texts, appear at the earliest stage. These are followed by Labeled Visual Texts, where basic textual features like labels are added to visuals, indicating an emerging awareness of multimodal meaning-making. Foundational Multimodal Texts mark a developmental shift, incorporating both written and visual elements with modest integration. Text-Centric Narrative Texts privilege written language but begin to show more organized structures and coherence. At more advanced levels, Redundant Complex Multimodal Texts show deliberate overlap between text and visuals to reinforce key ideas. Finally, Complementary Complex Multimodal Texts represent the highest level of hierarchical organization, where students strategically allocate different content across modes to maximize communicative clarity and epistemic function.

To support interpretation and comparison across levels, Table 6 summarizes the structural, modal, and content-based characteristics of each typology. It includes grade-level distribution, predominant text structures, dominant mode types, degree of mode integration, and content distribution patterns. This overview provides a visual hierarchy of the typologies and reinforces their placement along a developmental continuum of multimodal competency. The following sections elaborate on each typology in this sequence, highlighting their defining cognitive, structural, and communicative features.

TABLE 5 Latent class analysis fit indices and classification coefficients.

<i>k</i>	<i>d</i>	LL	BIC	SABIC	CAIC	AWE	BLRT <i>p</i>	VLMR-LRT <i>p</i>	Entropy	BF	<i>cmp</i>	% Largest class	% Smallest class
1	28	−40,746.17	81,700.69	81,611.74	81,610.82	81,624.82							
2	54	−38,638.57	77,678.97	77,507.42	77,505.65	77,532.65	0.00	0.00	0.89	0.000	0.00	61.4%	38.6%
3	80	−37,285.56	75,166.42	74,912.27	74,909.65	74,949.65	0.00	0.00	0.93	0.000	0.00	30.3%	14.9%
4	106	−37,127.23	75,043.23	74,706.48	74,703.01	74,756.01	0.00	0.00	0.94	0.002	0.00	34.1%	13.0%
5	132	−36,257.77	73,497.80	73,078.45	73,074.13	73,140.13	1.00	0.29	0.96	0.000	0.00	36.7%	6.0%
6	158	−35,743.23	72,662.18	72,160.23	72,155.06	72,234.06	0.00	0.76	0.94	0.000	1.00	21.9%	6.6%
7	184	−35,131.49	71,632.18	71,047.63	71,041.61	71,133.61	1.00	0.77	0.97			22.3%	0.1%

k, number of classes; *d*, number of parameters; LL, log-likelihood; BIC, Bayesian Information Criterion; SABIC, Sample-size adjusted BIC; CAIC, Consistent Akaike Information Criterion; AWE, Approximate Weight of Evidence Criterion; BLRT, bootstrapped likelihood ratio test; VLMR-LRT, Vuong-Lo-Mendell-Rubin adjusted likelihood ratio test; *p*, *p*-value; BF, Bayes Factor; *cmp*, correct model probability. Bolded values indicate “best” fit for each respective statistic.

TABLE 6 Overview of multimodal typologies by degree of hierarchical organization.

Typology	Grade-level distribution	Predominant text structures	Dominant mode use	Mode integration	Content distribution
Unlabeled visual texts	Mostly kindergarten	Description	Visuals only (mostly unlabeled illustrations)	Minimal	Visual only dominant; little text or overlap
Labeled visual texts	Mostly kindergarten	Description	Visuals with text labels (e.g., diagrams)	Emerging	Visual dominant with some text integration
Foundational multimodal texts	Mostly 1st grade	Description	Visuals with some linguistic features (e.g., titles)	Developing	Balanced mix of text, visual, and integrated
Text-centric narrative texts	1st and 2nd grade	Description and compare/contrast	Text-dominant (lists, emphasized text) with visuals	Moderate	Primarily text-based; minimal visual overlap
Redundant complex multimodal texts	Mostly 2nd grade	Compare/contrast and description	Text and visuals with significant duplication	Advanced	High overlap of text and visual content
Complementary complex multimodal texts	Mostly 2nd grade	Compare/contrast and description	Strategic use of text and visuals with minimal overlap	Highly Advanced	Visual dominant; distinct, complementary modes

Typologies are presented in order of increasing hierarchical organization. Mode integration reflects the degree to which linguistic and visual-graphical elements are used in coordinated, strategic ways. Content distribution refers to the proportion of information conveyed via text only, visuals only, or integrated across both.

Typology 1: complementary complex multimodal texts (N = 112 samples)

Rationale for Label: Named for their comprehensive integration of multiple modes, Complex Complementary Multimodal Texts were characterized by a balanced integration of modes, with a total content count of 10.6 units on average, and a significant emphasis on visual content (75.3%). The minimal overlap between text and visual content (1.4%) emphasized a strategic approach to multimodal integration and was predominantly observed in 2nd grade samples (97.3%).

Text Structure and Mode Frequency. The samples were primarily framed using compare/contrast (75.0%) and description (23.2%) structures, with an average of nearly 6 linguistic modes and 16.0 visual-graphical modes per sample. This indicated a cohesive use of these modes to organize and highlight information effectively.

Mode Types. Titles (83.0%), emphasized text (39.3%), and prefaces and conclusions (49.1%) were prevalent, indicating a strategic use of linguistic modes to organize and highlight information. Labeled illustrations (85.7%) and a variety of diagrams dominated the visuals, evidencing a rich engagement with visual content to enhance communication.

Content. Visual content was predominantly used (75.3%), followed by written text (23.4%), with minimal overlap between modes, suggesting a deliberate strategy to independently leverage the strengths of each mode for clear communication. Anecdotally, these samples featured written text with longer sentences but fewer scientific ideas. Visual content featured multiple simple labeled illustrations (e.g., if the text was about matter, samples contained multiple illustrations of various examples of matter like a glass of milk labeled milk). As indicated by the lack of duplication across mode, the visual content was rarely mentioned in the written text.

Typology 2: redundant complex multimodal texts (N = 373 samples)

Rationale for Label: These texts were distinguished by multiple redundancies across written text and visual content, with more than half of all content presented in both text and visual form (55.5%). These texts also contained the most content units (27.6) and were produced overwhelmingly by 2nd graders (75.3%).

Text Structure and Mode Frequency. Compare/contrast (64.3%) and descriptive (32.2%) structures were used to frame content, while samples contained an average of 6.3 linguistic modes and 13.3 visual-graphical modes per sample.

Mode Types. These samples featured a diverse set of linguistic modes including titles (68.1%), emphasized text (42.9%), headings (41.6%), and prefaces and conclusions (31.6%). A variety of visual-graphical modes were also used, notably labeled (74.0%) and unlabeled (41.3%) illustrations and tables (39.9%).

Content. These texts contained significant duplication of written text and visual content (55.5%) followed by visual only (24.2%) and textual only (16.9%).

Typology 3: foundational multimodal texts (N = 312 samples)

Rationale for Label: These texts demonstrated early multimodal engagement, with a total content count of 17.7 units. They were most commonly produced by 1st graders (93.9%).

Text Structure and Mode Frequency. These samples mostly used descriptions (89.1%) to frame their content, with an average of 1.9 linguistic modes and 12.4 visual-graphical modes per sample.

Mode Types. Frequent title use (47.1%) suggested an emerging understanding of textual organization. Labeled illustrations (13.8%) and a reliance on visual aids like labels (99.7%) and connectors (98.7%) complemented the textual information.

Content. These texts provided a balanced approach to content presentation, with visual (47.0%), written text (20.1%), and mixed text and visual content (32.9%), reflecting early attempts at multimodal communication.

Typology 4: unlabeled visual texts ($N = 313$ samples)

Rationale for Label. These texts were dominated by visual elements without textual labels, with a total content count of 10.8 units.

Text Structure and Mode Frequency. These samples largely used description structures (89.1%), with minimal use of linguistic modes (average of 0.23 per sample) while visual-graphical modes averaged 1.45 per sample.

Mode Types. There was minimal use of linguistic modes, with titles (10.5%) appearing most frequently. Visual-graphical modes were dominated by unlabeled illustrations (62.3%) indicating a primary reliance on visual representation.

Content. About half of the content was delivered via visuals (50.3%) with minimal combined text and visual content (24.3%) and text only content (14.5%), emphasizing visual descriptions.

Typology 5: text-centric narrative texts ($N = 253$ samples)

Rationale for Label. Defined by their reliance on written text to convey scientific concepts, these samples contained the lowest total content count (5.0) with the highest percentage of written text only content (96.0%). This typology was evenly distributed across 1st (48.6%) and 2nd grades (46.2%) with just 5.1% produced by kindergarteners.

Text Structure and Mode Frequency. Texts in this typology favored description (58.5%) and compare/contrast (40.3%) text structures, with frequent use of linguistic (average of 3.3 per sample) and visual-graphical modes (average of 4.5 per sample).

Mode Types. Titles (57.7%), lists (17.8%), and emphasized text (16.6%) were prevalent. There was limited use of visual-graphical modes with unlabeled (60.1%) and labeled illustrations (25.3%) and tables (19.0%) supplementing textual content.

Content. Nearly all content was delivered via written text only (96.0%), highlighting the typology's reliance on narrative descriptive texts for scientific communication.

Typology 6: labeled visual texts ($N = 342$ samples)

Rationale for Label: These texts were differentiated by the strategic use of labels to enhance visual representations, with a total content count of 17.9 units.

Text Structure and Mode Frequency. These texts were most frequently framed by description structures (91.8%) with minimal use of linguistic modes (0.12 per sample) and a significant use of visual-graphical modes (9.77 per sample).

Mode Types. These texts evidenced minimal engagement with linguistic modes. Titles (5.6%) were the most frequently found linguistic mode. Conversely, there was extensive use of labeled diagrams (83.0%) followed by illustrations (9.6%) and cycle diagrams (8.2%).

Content. These texts were comprised of visual (66.0%) or combined text and visual content (25.9%), reflecting strategic multimodal communication for clarity and comprehension. Anecdotally, these samples were overwhelmingly about plants. Typically, a flower was drawn along with the sun, dirt, and rain drops and labels for each, with this same visual content included in the written text (e.g., “a plant needs sun, water, and soil to grow”).

Discussion

Framed within a hierarchical model of scientific meaning-making progressing from discourse and genre to text structure, mode, and content, the analysis identified six distinct multimodal competency typologies using LCA: Complementary Complex Multimodal Texts, Redundant Complex Multimodal Texts, Foundational Multimodal Texts, Unlabeled Visual Texts, Text-Centric Narrative Texts, and Labeled Visual Texts. Rather than representing a fixed developmental sequence, these typologies illustrate the varied strategies that children employ to communicate scientific ideas and engage in the epistemic process of meaning-making through multimodal expression. When student texts are framed as epistemic tools (Prain and Hand, 2016) and writing is understood as a generative activity (Wittrock, 1992; Fiorella and Mayer, 2015), these typologies can be interpreted as manifestations of core cognitive processes including metacognition, selective integration, and an emerging understanding of mode affordances. Each typology captures a distinct approach through which young learners externalize, organize, and refine their ideas via the coordinated use of written language and visual-graphical modes.

Hierarchical organization in scientific communication

The framework of hierarchical organization provides a useful lens for understanding how students gradually develop the ability to integrate layers of scientific communication. At more advanced levels, learners begin to coordinate discourse conventions, genre expectations, text structures, representational modes, and scientific content into coherent, conceptually rich texts. This coordination enhances communicative clarity and facilitates conceptual elaboration, consistent with the epistemic function of multimodal composition (Ainsworth, 2006; Kress and van Leeuwen, 2021; Mayer, 2020). As students engage in scientific writing, their products become more than a simple record of knowledge. Even at early stages, children make meaningful decisions about how to select, organize, and transform ideas across representational forms (Lemke, 1998; Tang, 2020). Writing, labeling, diagramming, and sequencing thus become generative acts through which understanding is actively organized and revised.

Processes supporting hierarchical organization

The development of hierarchical organization in students' multimodal writing was supported by three interrelated processes: metacognition, selective integration, and an understanding of mode affordances. Together, these processes allow learners to construct texts that are both communicatively effective and epistemically generative. Metacognition refers to the ability to monitor, evaluate, and adjust one's thinking and learning strategies (Veenman et al., 2006). In the context of scientific writing, metacognitive awareness enables students to reflect on what they know, assess the effectiveness of their representational choices, and revise their texts for clarity and coherence. This includes decisions about how to organize information and select appropriate modes, often requiring simplification or emphasis, especially when composing for a younger peer, as in the current study, where simplification and emphasis are often required.

Selective integration involves the purposeful coordination of content and structure across multiple modes, supporting coherence within and across textual elements. Learners draw on text structures, linguistic elements, and visual-graphical features to produce cohesive, complementary representations of scientific ideas. Central to generative learning theory (Fiorella and Mayer, 2015), this process allows students to actively organize and transform content. Rather than duplicating information across modes, learners demonstrating greater multimodal competency make deliberate decisions about what is best conveyed through written language and what is more effectively illustrated through diagrams, tables, or labeled images.

An understanding of mode affordances, or the capacities and constraints of different representational systems, is also essential. From a cognitive perspective, students must learn to recognize which modes support particular conceptual demands (e.g., linguistic modes for abstract elaboration and temporal sequencing; visual-graphical modes for spatial relations, comparisons, or categorization; Ainsworth, 2006; Kress, 2010). However, affordances are not solely cognitive; they are also semiotic, shaped by the social, cultural, and disciplinary conventions that determine how meaning is made within specific contexts (Bezemer and Kress, 2008; Jewitt et al., 2016). Developing multimodal proficiency, therefore, involves recognizing what a mode can do cognitively as well as understanding how modes are used communicatively within a given genre or discipline. As students gain more experience with multimodal texts and instructional support, they move from surface-level inclusion of modes to more principled, epistemically grounded decisions about how to combine them effectively to convey scientific meaning.

Developmental patterns in multimodal competency

The six typologies identified in this study suggest a developmental continuum of representational practice as

well as increasing cognitive orchestration across multimodal representations. Lower-level typologies (i.e., Unlabeled and Labeled Visual Texts) rely heavily on listing or labeling with limited integration, suggesting early stages of idea encoding (Ainsworth, 2006; Tang et al., 2014). In contrast, higher-order types (i.e., Redundant and Complementary Complex Multimodal Texts) demonstrate coordination between modes, thematic organization, and abstraction, hallmarks of selective integration and metacognitive planning (Brod, 2021; Fiorella and Mayer, 2015). These more sophisticated compositions suggest a growing awareness of how to use different modes not just as decorative or illustrative elements (Carney and Levin, 2002; Fingeret, 2012), but as epistemic tools that support meaning-making, explanation, and reasoning (diSessa, 2004; Prain and Hand, 2016). Importantly, these shifts are not purely developmental in nature. Rather, they are mediated by children's understanding of genre and text structure, a form of cultural and communicative knowledge that provides scaffolding for more elaborate expression (Kress, 2010; McNeill and Krajcik, 2012).

Children may move fluidly between typologies depending on the topic at hand, the demands of the writing task, prior instructional experiences, and their comfort with different representational modes (Waldrip et al., 2010). This flexibility suggests that multimodal competency is not acquired in a linear fashion but instead develops as an adaptive skill set influenced by both cognitive growth and contextual support (Jewitt et al., 2016). What connects the typologies developmentally is a growing capacity to coordinate the layered components of scientific communication (i.e., discourse, genre, text structure, mode, and content) through increasingly deliberate and strategic cognitive processes (Mayer, 2020; Kress and van Leeuwen, 2021). As students develop metacognitive awareness, learn to selectively integrate information across modes, and gain experience with the affordances of visual and linguistic representations, their multimodal texts become more coherent, purposeful, and conceptually rich (Ainsworth, 2014; Schleppegrell, 2004). To illustrate developmental contrasts, the typologies are grouped into three broad clusters: Early Emergent, Transitional, and Developing Sophistication.

Early emergent typologies: unlabeled and labeled visual texts

Unlabeled Visual Texts, typically created by kindergarteners, consist of image-only compositions that rely on intuitive, concrete representations. Although they lack linguistic structure, these texts reflect an important epistemic function: allowing students to externalize and reflect on experience through visual means (Jewitt et al., 2016). Students at this level are beginning to engage with the mode and content layers of scientific writing but have not yet accessed genre or structure. Labeled Visual Texts represent an early step toward multimodal coordination. Students begin to pair single-word labels with visuals, indicating initial awareness of the affordances of language in reinforcing or clarifying meaning. These texts suggest emerging metacognitive insight into the limitations of visual-only communication, though the integration of modes remains minimal.

Transitional typologies: foundational multimodal and text-centric narrative texts

Foundational Multimodal Texts, most common in 1st grade, include simple descriptive sentences and accompanying visuals, although with loose or partial alignment. The use of text structure often in the form of basic descriptive or sequence formats signals that students are beginning to engage with the genre and structure layers of scientific discourse. At this stage, students are experimenting with selective integration, though the connections across modes are often underdeveloped. By contrast, Text-Centric Narrative Texts reflect stronger linguistic control and genre awareness but limited visual integration. Often produced by older students, these texts may reflect the broader emphasis on verbal modes in early literacy instruction, potentially at the expense of visual reasoning. While text structure is typically present and metacognitive control over language is evident, students may miss opportunities to use visuals as epistemic tools, thereby limiting the generative potential of their compositions (Schleppegrell, 2004; Waldrup et al., 2010).

Developing sophistication: redundant and complementary complex multimodal texts

Complex Redundant Multimodal Texts reflect an important transitional form of multimodal sophistication. These texts include a high volume of content, with more than half presented redundantly across text and visuals. This redundancy, while seemingly inefficient, may reflect a developmentally productive strategy: reinforcing meaning across modes to support conceptual consolidation (Schnotz and Bannert, 2003; Fiorella and Mayer, 2015). Students are beginning to recognize the value of visual-graphical representation but have not yet mastered strategic distribution of information. Still, their use of diverse linguistic and visual features demonstrates increased metacognitive control and an expanding awareness of mode affordances.

In contrast, Complex Complementary Multimodal Texts are marked by a strategic division of labor across modes, with minimal redundancy. These texts feature labeled diagrams, compare/contrast structures, and content distributed purposefully between written and visual elements. They reflect an emerging understanding that modes can serve complementary epistemic functions with visuals often used for classification or spatial relationships and text used for elaboration or comparison. However, these texts also featured fewer total content units and less diversity in mode types compared to Redundant texts. This suggests that while students producing Complementary texts may be making more deliberate decisions, they are still developing representational fluency, or the ability to flexibly combine a broad range of modes. Following Painter et al. (2013), these patterns may also reflect varying degrees of Commitment, where students' semiotic choices index differing levels of epistemic certainty, precision, or abstraction in how scientific knowledge is conveyed. These texts are promising indicators of sophistication, but not yet mastery. Collectively, these two typologies represent distinct expressions of multimodal growth. Redundant texts emphasize

breadth and reinforcement, while Complementary texts emphasize precision and differentiation. Both serve as productive sites for generative learning, though they reflect different stages in the coordination of representational and cognitive resources. This analysis also builds on Bateman's call for more integrated, empirically grounded studies of multimodal genre development, extending this work into early childhood science education where children begin to orchestrate semiotic resources to construct scientific meaning (Bateman, 2008; Bateman and Wildfeuer, 2014).

Implications for science education and early writing instruction

The present findings offer several implications for science instruction, writing pedagogy, and curriculum design in early education. First, scientific writing should be framed as both a form of communication and a process of knowledge generation. Teachers should model multimodal composition as a generative and epistemic practice, helping students externalize, organize, and transform their understanding across representational forms (Prain and Hand, 2016; Tang et al., 2014). Second, instructional scaffolding should support metacognition and strategic integration through on the use of just-in-time instruction. Just-in-time instruction is a form of explicit guidance that is provided at the moment it is needed, rather than through comprehensive instruction delivered upfront. This approach allows teachers to target support based on emerging student needs, helping avoid cognitive overload while still promoting strategic integration of content and modes (Ainsworth, 2014). This includes explicit instruction not only about the inclusion of images, but about how visual modes such as diagrams, labels, and comparison structures function as meaning-making resources in scientific contexts. Third, curriculum design should reflect the layered nature of scientific representation. Early writing instruction often emphasizes content and linguistic modes. However, helping students engage with discourse conventions, genre structures, and visual-graphical modes is critical for supporting deeper conceptual understanding and long-term engagement with science. Instructional design that reflects the layered framework used in this study (discourse → genre → structure → mode → content) can support more coherent development of multimodal scientific literacy (Fang, 2006; Kress, 2010).

Limitations and future directions

Although this study offers a detailed analysis of young children's multimodal scientific writing, several limitations should be considered. First, the data were collected from a rural region in the U.S. Midwest, which may limit the generalizability of findings to more diverse educational contexts. Future research should examine how multimodal competencies emerge across a broader range of socio-cultural and linguistic settings to better understand the role of contextual variation in shaping representational practices. Second, while students were asked to produce scientific summaries for a younger peer, this framing may have influenced the structure, tone, or modal choices of their texts. Further studies could examine how different rhetorical purposes or audiences shape students'

multimodal decisions. Third, although the study drew on a large dataset of writing samples from kindergarten through second grade, the data were analyzed cross-sectionally. As such, it does not capture individual developmental trajectories or instructional histories. Longitudinal research is needed to trace how students' multimodal competencies evolve over time and how instructional interventions may support progression through different typologies. Moreover, the study focused on students' written products rather than the compositional processes that produced them. While content analysis yields robust insights into representational outcomes, it does not illuminate the real-time decision-making that occurs during composition. Future process-oriented studies using think-aloud protocols, classroom video observations, or interviews could help uncover the cognitive and metacognitive strategies learners use as they select, integrate, and coordinate modes.

In addition, although all participating teachers received a shared professional development session, variability in classroom implementation, teacher expertise, and instructional time likely influenced student outputs. This demonstrates the need for further research exploring how specific instructional strategies including modeling, peer discussion, and explicit guidance in mode selection support the development of multimodal competency. Of particular interest is the role of just-in-time instruction, where explicit support is offered at the moment it is needed rather than delivered comprehensively upfront. This type of targeted scaffolding may reduce cognitive load while better supporting students' evolving understanding of how and when to use specific representational tools.

Finally, the content domains included in this study (plants, animals, and matter) may differ in how readily they afford visual or comparative representation. Future research should explore whether similar multimodal typologies emerge across other disciplinary areas or more abstract scientific content. It will also be valuable to examine how early multimodal competencies relate to later academic performance in science, literacy, and STEM fields more broadly. Taken together, these directions point to the importance of longitudinal, process-focused, and contextually sensitive research in understanding how multimodal scientific writing supports conceptual and disciplinary learning from the earliest years of schooling.

Conclusion

This study offers a developmental framework for understanding how young children engage in multimodal scientific communication. Drawing on over 1,700 student writing samples from kindergarten through 2nd grade, the analysis identified six distinct typologies that reflect students' evolving ability to coordinate structure, content, and mode within scientific texts. Rather than interpreting these typologies as fixed stages, the study conceptualizes multimodal competency as a flexible and adaptive system of representational strategies shaped by developmental readiness, instructional support, content familiarity, and prior experience. Crucially, these findings

contribute to a growing understanding of how cognition and multimodality intersect in young learners' scientific writing. When children engage in multimodal composition, they are doing more than transcribing what they know. They are actively integrating discourse expectations, genre conventions, and representational possibilities. This coordination requires the kinds of cognitive moves emphasized by generative learning theory, especially metacognition, abstraction, and integration, revealing the epistemic function of multimodal artifacts. As such, multimodal texts function as epistemic tools, enabling learners to externalize, refine, and reconfigure their understanding. By identifying patterns of representational practice and connecting them to key cognitive processes, this study provides new evidence regarding how young learners build multimodal fluency in science. Supporting the development of these competencies is central to effective communication as well as essential to cultivating the foundational skills that enable children to think, reason, and participate as scientific learners. Continued research and intentional instructional design will be critical for advancing both the theory and practice of multimodal scientific literacy in the early years.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by Purdue University Institutional Review Board. The studies were conducted in accordance with the local legislation and institutional requirements. The Ethics Committee/Institutional Review Board waived the requirement of written informed consent for participation from the participants or the participants' legal guardians/next of kin because The summary writing samples were provided to the lead author in a de-identified digital format according to a data sharing agreement between Purdue University and the University of Iowa.

Author contributions

DN: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. TN: Conceptualization, Investigation, Validation, Writing – original draft, Writing – review & editing. BH: Conceptualization, Data curation, Investigation, Methodology, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. The lead author received

a grant from the Caplan Foundation for Early Childhood to support this research.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

References

- Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learn. Instruct.* 16, 183–198. doi: 10.1016/j.learninstruc.2006.03.001
- Ainsworth, S. (2014). “The multiple representations principle in multimedia learning,” in *The Cambridge Handbook of Multimedia Learning*, ed. R. E. Mayer (Cambridge: Cambridge University Press), 158–170.
- Ainsworth, S., Tytler, R., and Prain, V. (2020). “Learning by construction of multiple representations,” in *Handbook of Learning from Multiple Representations and Perspectives*, eds. P. Van Meter, A. List, D. Lombardi, and P. Kendeou (New York, NY: Routledge), 92–106.
- Bartlett, B. J. (1978). *Top-Level Structure as an Organizational Strategy for Recall of Classroom Text* (Order No. 7911113) (Available from ProQuest Dissertations and Theses Global Closed Collection; ProQuest One Academic). Available at: <https://www.proquest.com/dissertations-theses/top-level-structure-as-organizational-strategy/docview/302878589/se-2>
- Bateman, J. (2008). *Multimodality and Genre: A Foundation for the Systematic Analysis of Multimodal Documents*. Berlin: Springer.
- Bateman, J., Wildfeuer, J., and Hiippala, T. (2017). *Multimodality: Foundations, Research and Analysis—A Problem-Oriented Introduction*. Berlin: de Gruyter Mouton.
- Bateman, J. A., and Wildfeuer, J. (2014). A multimodal discourse theory of visual narrative. *J. Pragmat.* 74, 180–208. doi: 10.1016/j.pragma.2014.10.001
- Beck, C. R. (1984). Visual cueing strategies: pictorial, textual and combinational effects. *Educ. Technol. Res. Dev.* 32, 207–216. doi: 10.1007/BF02768892
- Bezemer, J., and Kress, G. (2008). Writing in multimodal texts: A social semiotic account of designs for learning. *Written Commun.* 25, 166–195. doi: 10.1177/0741088307313
- Bilalić, M., McLeod, P., and Gobet, F. (2010). The mechanism of the Einstellung (set) effect: a pervasive source of cognitive bias. *Curr. Dir. Psychol. Sci.* 19, 111–115. doi: 10.1177/0963721410363571
- Boucheix, J., Lowe, R., and Thibaut, J. (2020). A developmental perspective on young children's understandings of paired graphics conventions from an analogy task. *Front. Psychol.* 11:2032. doi: 10.3389/fpsyg.2020.02032
- Bovair, S., and Kieras, D. E. (1991). “Toward a model of acquiring procedures from text,” in *Handbook of Reading Research, Vol II*, eds. R. Barr, M. L. Kamil, P. B. Mosenthal, and P. D. Pearson (Mahwah: Erlbaum), 206–229.
- Bransford, J. D., Brown, A. L., and Cocking, R. R. (2000). *How People Learn: Brain, Mind, Experience, and School: Expanded edition*. Washington, DC: National Academy Press.
- Brod, G. (2021). Generative learning: which strategies for what age? *Educ. Psychol. Rev.* 33, 1295–1318. doi: 10.1007/s10648-020-09571-9
- Carney, R. N., and Levin, J. R. (2002). Pictorial illustrations still improve students' learning from text. *Educ. Psychol. Rev.* 14, 5–26. doi: 10.1023/A:1013176309260
- Carroll, J. B. (1993). *Human Cognitive Abilities: A Survey of Factor-Analytic Studies (No. 1)*. Cambridge: Cambridge University Press.
- Chen, H., Myhill, D., and Lewis, H. (Eds.). (2020). *Developing Writers Across the Primary and Secondary Years: Growing into Writing*. New York, NY: Routledge.
- Chi, M. T. H., Feltovich, P. J., and Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cogn. Sci.* 5, 121–152. doi: 10.1207/s15516709cog0502_2
- Christie, F., and Derewianka, B. (2008). *School Discourse: Learning to Write Across the Years of Schooling*. Bloomsbury: Bloomsbury Publishing.
- Davis, O. L., and Hunkins, F. P. (1966). Textbook questions: what thinking processes do they foster? *Peabody J. Educ.* 43, 285–292. doi: 10.1080/01619566609537358
- diSessa, A. A. (2004). Metarepresentation: Native competence and targets for instruction. *Cogn. Instr.* 22, 293–331. doi: 10.1207/s1532690xci2203_2
- Duke, N. K. (2000). 3.6 minutes per day: The scarcity of informational texts in first grade. *Read. Res. Q.* 35, 202–224. doi: 10.1598/RRQ.35.2.1
- Duke, N. K., and Tower, C. (2004). “Nonfiction texts for young readers,” in *The Texts in Elementary Classrooms*, eds. J. Hoffman, D. L. Schallert (Mahwah, NJ: Erlbaum), 111–128.
- Duke, N. K., Ward, A. E., and Pearson, P. D. (2021). The science of reading comprehension. *Read. Teach.* 74, 663–672. doi: 10.1002/trtr.1993
- Fang, Z. (2006). The language demands of science reading in middle school. *Int. J. Sci. Educ.* 5, 491–520. doi: 10.1080/09500690500339092
- Feltovich, P. J., Prietula, M. J., and Ericsson, K. A. (2006). “Studies of expertise from psychological perspectives,” in *The Cambridge Handbook of Expertise and Expert Performance*, eds. K. A. Ericsson, N. Charness, P. J. Feltovich, and R. R. Hoffman (Cambridge: Cambridge University Press), 41–67.
- Fingeret, L. (2012). *Graphics in children's informational texts: A content analysis* (Order No. 3524408) (Available from Dissertations and Theses @ Big Ten Academic Alliance; ProQuest Dissertations and Theses Global Closed Collection; ProQuest One Academic) (1039317370). Available at: <https://www.proquest.com/dissertations-theses/graphics-childrens-informational-texts-content/docview/1039317370/se-2>
- Fiorella, L., and Mayer, R. E. (2015). *Learning as a Generative Activity*. Cambridge: Cambridge University Press.
- Fisher, D., Frey, N., and Lapp, D. (2012). Building and activating students' background knowledge: It's what they already know that counts: teachers must assess and build on the background knowledge students possess. *Middle School J.* 43, 22–31. doi: 10.1080/00940771.2012.11461808
- Flavell, J. H. (1979). Metacognition and cognitive monitoring: a new area of cognitive-developmental inquiry. *Am. Psychol.* 34, 906–911. doi: 10.1037/0003-066X.34.10.906
- Franconeri, S., Padilla, L., Shah, P., Zacks, J. M., and Hullman, J. (2021). The science of visual data communication: what works. *Psychol. Sci. Public Interest* 22, 110–161. doi: 10.1177/15291006211051956
- Glushko, R. J. (2013). *The Discipline of Organizing: 4th Professional Edition*. Berkeley, CA: UC Berkeley.
- Gobet, F., Lane, P. C., Croker, S., Cheng, P. C., Jones, G., Oliver, I., et al. (2001). Chunking mechanisms in human learning. *Trends Cogn. Sci.* 5, 236–243. doi: 10.1016/S1364-6613(00)01662-4

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcomm.2025.1603737/full#supplementary-material>

- Goldman, S. R., and Rakestraw, J. A. (2000). "Structural aspects of constructing meaning from text," in *Handbook of Reading Research, Vol. III*, eds. M. L. Kamil, P. B. Mosenthal, P. D. Pearson, and R. Barr (Mahwah, NJ: Erlbaum), 311–335.
- Graham, S., Kiuhara, S. A., and MacKay, M. (2020). The effects of writing on learning in science, social studies, and mathematics: a meta-analysis. *Rev. Educ. Res.* 90, 179–226. doi: 10.1012/0034654320914744
- Guo, D., Wright, K. L., and McTigue, E. M. (2018). A content analysis of visuals in elementary school textbooks. *Elem. Sch. J.* 119, 244–269. doi: 10.1086/700266
- Halford, G. S. (2014). *Children's Understanding: The Development of Mental Models*. New York, NY: Psychology Press.
- Halliday, M. A. K., and Martin, J. R. (1993). *Writing Science: Literacy and Discursive Power*. New York, NY: Routledge.
- Hand, B. (2008). "The case for the SWH approach," in *Science Inquiry Argument and Language: A Case for the Science Writing Heuristic* (Rotterdam: Sense Publishers), 195–205.
- Hand, B., Chen, Y. C., and Suh, J. K. (2021). Does a knowledge generation approach to learning benefit students? A systematic review of research on the Science Writing Heuristic approach. *Educ. Psychol. Rev.* 33, 535–577. doi: 10.1007/s10648-020-09550-0
- Harp, S. F., and Mayer, R. E. (1998). How seductive details do their damage: a theory of cognitive interest in science learning. *J. Educ. Psychol.* 90, 414–434. doi: 10.1037/0022-0663.90.3.414
- Hartley, J. (2014). Current findings from research on structured abstracts: an update [comment and opinion]. *J. Med. Libr. Assoc.* 102, 146–148. doi: 10.3163/1536-5050.102.3.002
- Hess, K. (2010). *Teaching and Assessing Understanding of Text Structures Across Grades*. Center for Assessment. Available at: <https://www.nciea.org/library/teaching-and-assessing-understanding-of-text-structures-across-grades/> (Retrieved September 20, 2023).
- Humphrey, S. (2018). "We Can Speak to the World": applying meta-linguistic knowledge for specialized and reflexive literacies," in *Bilingual Learners and Social Equity. Educational Linguistics, Vol. 33*, ed. R. Harman (Cham: Springer). doi: 10.1007/978-3-319-60953-9_3
- Hyland, K. (2015). Genre, discipline and identity. *J. English Acad. Purp.* 19, 32–43. doi: 10.1016/j.jeap.2015.02.005
- Jewitt, C., Bezemer, J., and O'Halloran, K. (2016). *Introducing Multimodality*. New York, NY: Routledge.
- Kintsch, W. (1998). *Comprehension: A Paradigm for Cognition*. Cambridge: Cambridge University Press.
- Kobayashi, M. (2002). Method effects on reading comprehension test performance: text organization and response format. *Lang. Test.* 19, 193–220. doi: 10.1191/0265532202lt2270a
- Kress, G. (2010). *Multimodality: A Social Semiotic Approach to Contemporary Communication*. New York, NY: Routledge. Available at: <http://ci.ni.ac.jp/ncid/B800321027>
- Kress, G., and van Leeuwen, T. (2021). *Reading images: The grammar of visual design (3rd Ed.)*. New York, NY: Routledge. doi: 10.4324/9781003099857
- Krippendorff, K. (2019). *Content Analysis: An Introduction to its Methodology (4th Ed.)*. Thousand Oaks: Sage.
- Lamberski, R. J., and Dwyer, F. M. (1983). The instructional effect of coding (color and black and white) on information acquisition and retrieval. *Educ. Technol. Res. Dev.* 31, 9–21. doi: 10.1007/BF02765207
- Landis, J. R., and Koch, G. G. (1977). An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics* 33, 363–374. doi: 10.2307/2529786
- Lemke, J. L. (1998). "Multiplying meaning: visual and verbal semiotics in scientific text," in *Reading Science*, eds. J. R. Martin and R. Veel (New York, NY: Routledge), 87–113.
- Lemke, J. L. (2004). "The literacies of science," in *Crossing Borders in Literacy and Science Instruction: Perspectives on Theory and Practice*, ed. E. W. Saul (Newark: International Reading Association), 33–47.
- Martin, J. R., and Rose, D. (2012). Genres and texts: living in the real world. *Indones. J. Syst. Funct. Linguist.* 1, 1–21.
- Masyn, K. E. (2013). "Latent class analysis and finite mixture modeling," in *The Oxford Handbook of Quantitative Methods: Statistical Analysis*, ed. T. D. Little (Oxford: Oxford University Press), 55–611.
- Mayer, R. E. (2020). *Multimedia Learning (3rd ed.)*. Cambridge: Cambridge University Press.
- McNeill, K. L., and Krajcik, J. (2012). *Supporting Grade 5–8 Students in Constructing Explanations in Science: The Claim, Evidence, and Reasoning Framework for Talk and Writing (1st Ed.)*. London: Pearson Education.
- Meyer, B. J. F. (1975). Identification of the structure of prose and its implications for the study of reading and memory. *J. Read. Behav.* 7, 7–47. doi: 10.1080/10862967509547120
- Meyer, B. J. F., Brandt, D. M., and Bluth, G. J. (1980). Use of top-level structure in text: Key for reading comprehension of ninth-grade students. *Read. Res. Q.* 16, 72–103. doi: 10.2307/747349
- Muthén, B., and Muthén, L. K. (2000). Integrating person-centered and variable-centered analyses: growth mixture modeling with latent trajectory classes. *Alcohol. Clin. Exp. Res.* 24, 882–891. doi: 10.1111/j.1530-0277.2000.tb02070.x
- NAEP (2022). "What does the NAEP reading assessment measure?" in National Center for Education Statistics. Available at: <https://nces.ed.gov/nationsreportcard/reading/whatmeasure.aspx> (Accessed January 23, 2023).
- Nagin, D. S. (2014). Group-based trajectory modeling: an overview. *Ann. Nutr. Metabol.* 65, 205–210. doi: 10.1159/000360229
- National Research Council (NRC) (2012). *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. Washington, DC: National Academies Press.
- NGSS Lead States (2013). *Next Generation Science Standards: For states, by states*. Washington, DC: The National Academies Press.
- Norris, S. P., and Phillips, L. M. (2003). How literacy in its fundamental sense is central to scientific literacy. *Sci. Educ.* 87, 224–240. doi: 10.1002/sc.10066
- Nylund-Gibson, K., and Choi, A. Y. (2018). Ten frequently asked questions about latent class analysis. *Transl. Issues Psychol. Sci.* 4, 440–461. doi: 10.1037/tps0000176
- Osborne, R., and Wittrock, M. (1985). The generative learning model and its implications for science education. *Stud. Sci. Educ.* 12, 59–87. doi: 10.1080/03057268508559923
- Painter, C., Martin, J. R., and Unsworth, L. (2013) *Reading Visual Narratives: Image Analysis of Children's Picture Books*. Sheffield: Equinox.
- Palincsar, A. S., Magnusson, S. J., Collins, K. M., and Cutter, J. (2001). Making science accessible to all: results of a design experiment in inclusive classrooms. *Learn. Disabil. Q.* 24, 15–32. doi: 10.2307/1511293
- Pettersson, R. (2019). "It depends," in *Institutet for Infology*. Available at: <https://www.iitd.net/PublicLibrary/Pettersson-Rune-ID-It-Depends.pdf> (Accessed March 6, 2023).
- Pilegard, C., and Fiorella, L. (2016). Helping students help themselves: Generative learning strategies improve middle school students' self-regulation in a cognitive tutor. *Comput. Human Behav.* 65, 121–126. doi: 10.1016/j.chb.2016.08.020
- Prain, V., and Hand, B. (2016). "Learning science through learning to use its languages," in *Using Multimodal Representations to Support Learning in the Science Classroom*, eds. B. Hand, M. McDermott, and V. Prain (Berlin: Springer).
- Prain, V., and Tytler, R. (2022). Theorising learning in science through integrating multimodal representations. *Res. Sci. Educ.* 52, 805–817. doi: 10.1007/s11165-021-10025-7
- Purcell-Gates, V., Duke, N. K., and Martineau, J. A. (2007). Learning to read and write genre-specific text: roles of authentic experience and explicit teaching. *Read. Res. Q.* 42, 8–45. doi: 10.1598/RRQ.42.1.1
- Roberts, K. L., Norman, R. R., Duke, N. K., Morsink, P. M., Martin, N. M., Knight, J. F., et al. (2013). Diagrams, timelines, and tables-oh, my! Fostering graphical literacy. *Read. Teach.* 67, 12–24. doi: 10.1002/TRTR.1174
- Schleppegrell, M. J. (2004). *The Language of Schooling: A Functional Linguistics Perspective*. New York, NY: Routledge.
- Schnotz, W., and Bannert, M. (2003). Construction and interference in learning from multiple representation. *Learn. Instr.* 13, 141–156. doi: 10.1016/S0959-4752(02)00017-8
- Skalski, P. D., Neuendorf, K. A., and Cajigas, J. A. (2019). *The Content Analysis Guidebook*. Thousand Oaks: Sage.
- Slutsky, D. J. (2014). The effective use of graphs. *J. Wrist Surg.* 03, 067–068. doi: 10.1055/s-0034-1375704
- Stöckl, H. (2004). "In between modes: language and image in printed media," in *Perspectives on Multi-Modality*, eds. E. Ventola, C. Charles, and M. Kaltenbacher (Amsterdam: John Benjamins Publishing Company), 9–30.
- Stöckl, H. (2020). "Linguistic multimodality—Multimodal Linguistics: a state-of-the-art sketch," in *Multimodality: Disciplinary Thoughts and the Challenge of Diversity*, eds. J. Wildfeuer, J. Pflaeging, J. Bateman, O. Seizov and C. Tseng (Berlin: De Gruyter), 41–68.
- Swales, J. M. (1990). *Genre Analysis: English in Academic and Research Settings*. Cambridge: Cambridge University Press.
- Sweller, J. (2011). "Cognitive load theory," in *The Psychology of Learning and Motivation: Cognition in Education*, eds. J. P. Mestre and B. H. Ross (Cambridge: Elsevier Academic Press), 37–76.
- Tang, K. S. (2020). The use of epistemic tools to facilitate epistemic cognition and metacognition in developing scientific explanation. *Cogn. Instr.* 38, 474–502. doi: 10.1080/07370008.2020.1745803
- Tang, K. S., Delgado, C., and Moje, E. B. (2014). An integrative framework for the analysis of multiple and multimodal representations for meaning-making in science education. *Sci. Educ.* 98, 305–326. doi: 10.1002/sc.21099

- Tippett, C. D. (2010). Refutation text in science education: a review of two decades of research. *Int. J. Sci. Math. Educ.* 8, 951–970. doi: 10.1007/s10763-010-9203-x
- Unsworth, L. (2001). *Teaching Multiliteracies Across the Curriculum: Changing Contexts of Text and Image in Classroom Practice*. London: Open University Press.
- van den Broek, P. (2010). Using texts in science education: Cognitive processes and knowledge representation. *Science* 328, 453–456. doi: 10.1126/science.1182594
- van Leeuwen, T. (2005). *Introducing Social Semiotics: An Introductory Textbook*. New York, NY: Routledge.
- van Merriënboer, J. J., and Sweller, J. (2010). Cognitive load theory in health professional education: design principles and strategies. *Med. Educ.* 44, 85–93. doi: 10.1111/j.1365-2923.2009.03498.x
- Van Rooy, W. S., and Chan, E. (2017). Multimodal representations in senior biology assessments: a case study of NSW Australia. *Int. J. Sci. Math. Educ.* 15, 1237–1256. doi: 10.1007/s10763-016-9741-y
- Veenman, M. V., Van Hout-Wolters, B. H., and Afflerbach, P. (2006). Metacognition and learning: conceptual and methodological considerations. *Metacogn. Learn.* 1, 3–14. doi: 10.1007/s11409-006-6893-0
- Waldrip, B., Prain, V., and Carolan, J. (2010). Using multi-modal representations to improve learning in junior secondary science. *Res. Sci. Educ.* 40, 65–80. doi: 10.1007/s11165-009-9157-6
- Wexler, N. (2020). *The Knowledge Gap: The Hidden Cause of America's Broken Education System—and How to Fix It*. Westminster: Penguin.
- Williams, J. P. (2018). Text structure instruction: the research is moving forward. *Read. Writ.* 31, 1923–1935. doi: 10.1007/s11145-018-9909-7
- Wittrock, M. C. (1974). Learning as a generative process. *Educ. Psychol.* 11, 87–95. doi: 10.1080/00461527409529129
- Wittrock, M. C. (1992). Generative learning processes of the brain. *Educ. Psychol.* 27, 531–541. doi: 10.1207/s15326985ep2704_8
- Wylie, R., and Chi, M. (2014). “The self-explanation principle in multimedia learning,” in *The Cambridge Handbook of Multimedia Learning*, ed. R. Mayer (Cambridge Handbooks in Psychology) (Cambridge: Cambridge University Press), 413–432.
- Yore, L. D., Bisanz, G. L., and Hand, B. M. (2003). Examining the literacy component of science literacy: 25 years of language arts and science research. *Int. J. Sci. Educ.* 25, 689–725. doi: 10.1080/09500690305018