Narrative Strategies in Picture Books:
The Role of Words, Illustrations, and their Interplay

Todsapon Suranakkharin
Faculty of Humanities, Naresuan University
Corresponding author: todsapons@nu.ac.th

Abstract

In this study, Nikolajeva and Scott's (2001) typology of text-image interactions is employed to analyze the narrative strategies in three selected picture books, including *Two Bad Ants* by Chris Van Allsburg, *Dear Diary* by Sara Fanelli, and *Sector 7* by David Wiesner. The results demonstrate that complementary and counterpoint interactions are predominantly used in *Two Bad Ants*, where text provides temporal and emotional depth while illustrations offer spatial clarity and work in transforming ordinary objects into life-threatening challenges. In contrast, a text-led approach is used to convey meaning in *Dear Diary*, where illustrations work to enhance humor and tone. *Sector 7,* on the other hand, relies entirely on visual storytelling, and thus offers readers interpretive freedom and opportunities to develop visual literacy. By contrasting various narrative strategies employed in these three selected picture books, this study demonstrates how different text-image interactions may engage readers, form meaning, and produce multi-layered narratives. Beyond literary analysis, these findings offer pedagogical and creative significance. They suggest ways for educators to use picture books to foster both verbal and visual literacy, and for picture book creators to apply diverse narrative strategies to enrich storytelling for different audiences. The study thus underscores the multifaceted potential of picture books as a multimodal narrative form that engages readers and supports meaningful literacy development.

Keywords: Text-image interaction, Narrative strategies, Multimodal storytelling, Picture books

Introduction

Picture books are profusely illustrated books in which both words and illustrations contribute to the story's meaning. Illustrations in picture books are integral to the story, providing actual plot or concept information as well as clues to character traits, settings, and moods (Lynch-Brown & Tomlinson, 2002, p. 76).

This observation by Lynch-Brown and Tomlinson reflects the inherent complexity of picture books. To most people, picture books are understood as comprising two principal elements: words and illustrations. When discussing picture books, it is clear that we are not focusing on words or illustrations in isolation but rather on the complex interplay between these two components (See, for example, Nodelman, 1988; Stephen, 1992; Nikolajeva & Scott, 2001; Lewis, 2001).

In his seminal work *Words About Pictures: The Narrative Art of Children's Picture Books*, Nodelman (1988) argues that the interaction between words and illustrations in picture books is far from straightforward. Rather than simply mirroring each other, these two components collaborate to communicate meaning through various narrative strategies. Typically, words work in providing temporal progression and emotional depth, while illustrations provide spatial clarity and visual perspective, allowing readers to perceive narrative elements that words alone cannot fully communicate. According to Nodelman (1988), illustrations that appear without accompanying words invite multiple interpretations, implying the openness and ambiguity that purely visual storytelling often brings. Conversely, without supporting illustrations, words may restrict the reader's imaginative engagement because of its absence of specific visual features. It is therefore the interplay between words and illustrations that creates a dynamic tension and enhances the depth and interpretive potential of picture books.

Based on the above assumption, there are always three distinct narrative layers delivered in picture books: one conveyed by words, one through illustrations, and another created by interplay between the two components (Arizpe & Styles, 2016; Pantaleo, 2008; Sipe, 2012). However, the text-image interactions vary according to each book's narrative structure and artistic choices, which significantly demonstrates how different strategies may shape storytelling in unique way. To further explore these dynamics, this study examines how varying narrative approaches and text-image relationships operate in three carefully selected picture books that exemplify different narrative approaches: Chris Van Allsburg's *Two Bad Ants*, Sara Fanelli's *Dear Diary*, and David Wiesner's *Sector 7*. These books were purposively selected based on three criteria: (1) each represents a distinctive type of text-image relationship ranging from text-dominant to image-dominant storytelling; (2) all are critically acclaimed works recognized for their innovative visual and narrative design; and (3) together they provide a balanced sample for examining the spectrum of multimodal interactions in contemporary picture books. With their distinct approaches to the text-image interplay, *Two Bad Ants* demonstrates a close balance and interdependence between verbal and visual storytelling; *Dear Diary* represents a text-led narrative enriched by illustrations that enhance tone and humor; while *Sector 7* exemplifies a purely visual mode of storytelling. By limiting the analysis to these three carefully selected books, it is also possible to offer an in-depth exploration of each picture book's narrative strategies, including how they engage readers, encourage interpretive freedom, and enhance visual literacy.

The Interplay Between Words and Illustrations

The analysis of selected picture books in this study draws on the typology proposed by Nikolajeva and Scott (2001), who emphasize the need for "a consistent and flexible terminology, a comprehensive international metalanguage, and a system of categories describing the variety of text/image interactions" in picture books (p. 6). In their book *How Picturebooks Work*, they provide a fine differentiation between types of text-image interactions, ranging from symmetry, through complementary, enhancement, counterpoint, and contradiction. They also explore how these categories interact with literary elements such as character, setting, and point of view, demonstrating the intricate ways in which picture books function as multimodal narratives.

To begin with, in symmetrical interaction, words and illustrations deliver identical information. This alignment reinforces clarity and ensures consistency, making it easier for readers to follow the story. This approach is particularly suitable for younger audiences or straightforward narratives. However, it can limit interpretive depth because it leaves little opportunity for readers to explore alternative perspectives. On the contrary, in a complementary interaction, different pieces of information are conveyed through text and illustrations. Each of these two components largely relies on the other to bridge gaps and collectively provide a cohesive narrative. Only by integrating the two modalities can readers achieve a deeper comprehension of the narrative, thereby enriching their overall reading experience.

Another type of text-image interaction is called enhancement, which focuses on how the two components may work in amplifying the emotional or thematic depth of the story. Visual elements, for instance, may augment humor, heighten suspense, or evoke poignancy. This can occur even in cases where words and illustrations do not introduce new or contrasting details. Conversely, counterpoint interaction creates complexity by opposing or conflicting information presented between text and images. This dynamic generates suspense and motivates readers to closely study the interaction between the two components. Often presenting uncertainty, counterpoint interaction calls for readers to actively interact imaginatively with the text; hence, it is especially successful in advanced picture books encouraging interpretive thinking.

Contradictory is the last and the most unconventional type in this framework. This interaction occurs when words and illustrations directly oppose each other, and readers are deliberately challenged to reconcile opposing information and construct their meaning. Although rare, this type of interaction fosters critical thinking and active engagement. It also evokes strong emotional responses and draws attention to themes such as irony, deception, or ambiguity. This framework forms the foundation for analyzing how text and illustrations in picture books interact. By applying these categories, this study investigates how these interactions influence narrative meaning and engage readers on both emotional and cognitive levels.

Tiny Adventurers: The Collaborative Narrative of Words and Illustrations in *Two Bad Ants*

Two Bad Ants has long been successful in captivating readers' attention with its engaging and adventurous plot. The story centers on two ants who decide not to return to their colony and instead stay behind to savor the reward alone. However, as their life outside the colony turns into a terrifying adventure, they eventually come to understand the value of community. The success of *Two Bad Ants*

relies heavily on the way the words and illustrations work collaboratively to convey meaning. To understand this interplay, it is essential to examine the distinct roles played by each component..

When considered independently, words are found to work well in providing vivid descriptions, but they clearly lack explicit spatial clarity. For example, when the ants set out the journey to collect more "crystals", they are described traveling through "the dark forest" and climbing "a mountain". These phrases gain specific significance only when coupled with visuals. The illustrations reveal that "crystals" actually refer to sugar, "the dark forest" is grass, and "a mountain" is a brick wall. This dynamic illustrates a complementary interaction, where text and illustrations provide unique yet interrelated information to create a unified narrative. With this interplay, readers are prompted to reimagine familiar settings and perceive the world through the ants' perspective (See Figure. 1) Here, the illustrations do not contradict the text, but they clarify and expand its meaning, enabling readers to understand the ants' journey on both the literal and figurative scales.

Figure.1 Two Bad Ants
(Source: Goss, 2020)

Conversely, illustrations alone cannot convey a unified story. Although cohesively arranged, illustrations lack details necessary to effectively communicate the story's intent. For example, the ants' journey and the crystals' significance are only clear when the text states, "They were eager to gather more crystals because the queen was the mother of them all". In other words, the text provides the motivation and context that the illustrations fail to provide. Similarly, temporal information is more efficiently communicated through time markers employed in the text: "It was late in the day when they departed… dusk turned to twilight, twilight to night". Without words, the passage of time become rather ambiguous since the visuals alone cannot show the journey's duration. This balance between these two elements underscores the interdependence of multimodal narratives, with words and illustrations filling gaps left by the other.

As the story progresses, the complementary interaction begins to shift toward counterpoint in the depiction of the ants' size in relation to the surroundings. While the ants' journey is textually described

as their routine exploration, their insignificance and vulnerability are visually emphasized. At the beginning of the story, when the ants are still in their tunnel, the ants are depicted as large and purposeful (See Figure.2). In sharp contrast, once they enter the outside world, the ants are portrayed as tiny figures against towering objects. At this juncture, a counterpoint dynamic is introduced. Instead of simply complementing the text by filling in missing details, the illustrations of ants in small size begin to reframe the narrative by highlighting the overwhelming scale of challenges the ants face. With this interplay, the ants' journey appears to be more terrifying than the words alone suggest. Readers are thereby drawn into the ants' vulnerable world, fostering empathy and deeper emotional engagement.
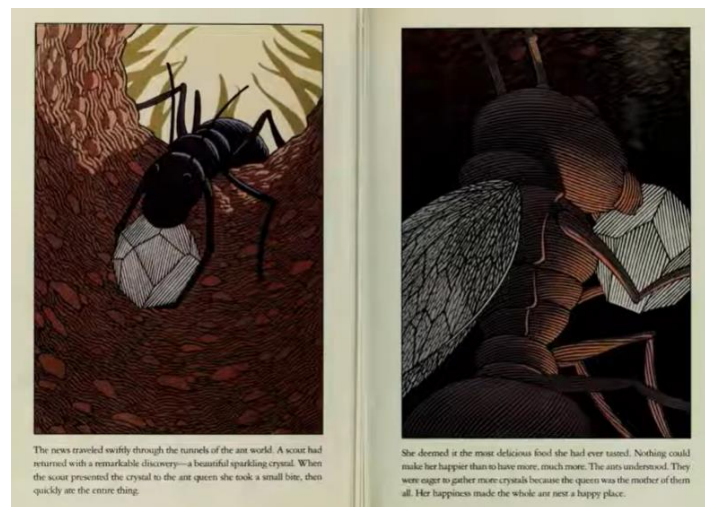

Figure.2 Two Bad Ants
(Source: Goss, 2020)

After the two ants decide to stay behind, ordinary household items are textually transformed into a series of life-threatening challenges. Each is described in dramatic terms from the ants' perspective: "a boiling brown lake", "a dark space with a strange red glow", "a waterfall", "a whirling storm in a dark chamber", and "narrow holes [with] dark openings". Here, the illustrations once again work effectively in a complementary interaction with the text by providing visual clues that those terrifying obstacles are, in fact, a cup of coffee, a toaster, a sink, a garbage disposal, and an electrical socket. The illustrations further compliment the text by making the imminent danger visually explicit and the two ants appear more powerless. For example, in the coffee scene, "a whirlpool that sucked the ants deeper and deeper" is visualized as a vast, overwhelming wave sweeping across the pages, with the ants barely visible. However, a counterpoint interaction also coexists in certain kitchen scenes, with the illustrations going beyond complimenting the text. For example, upon seeing the narrow holes, the text describes how the two ants are reminded of their "warmth and safety of their old underground home" before deciding to climb inside. The accompanying illustration, on the other hand, reveal that the ants are about to enter an electoral socket. This discrepancy between the ants' naivety the text and the reader's understanding conveyed by the illustration creates a counterpoint interaction. In other words, a moment of dramatic irony is created: the ants are calmly curious, but the readers see they are on the verge of a dangerous

mistake (See Figure.3). When combined, these complementary and counterpoint dynamics enhance the text-image relationship within the story. While complementary ensure narrative coherence by aligning text and illustrations, counterpoint interaction goes beyond the text and prompt readers to engage more critically and reinterpret the ants' experiences through imaginative inference.
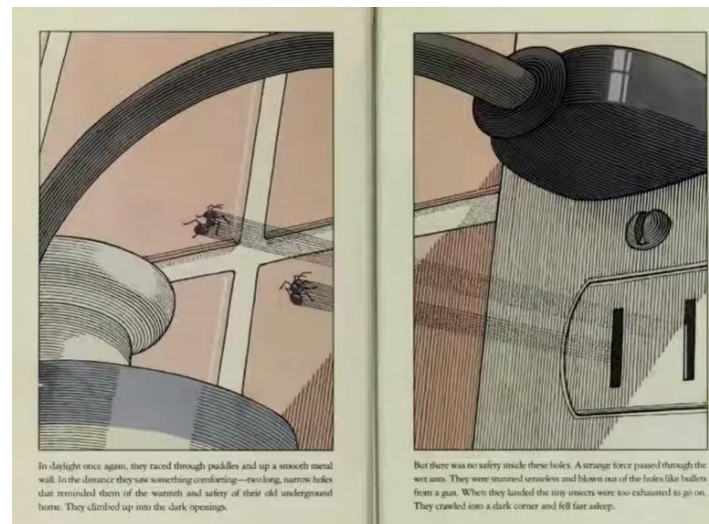

Figure.3 Two Bad Ants
(Source: Goss, 2020)

Towards the end of the story, complementary interaction is employed to convey the emotional impact of the ants' journey. After surviving from a series of life-threatening challenges, the two ants quietly rejoin the line of their freinds who return to collect more crystals and follow them home. In this scene, the illustration shows the two ants, each holding a single crystal, approach the entrance of the ant hole. Although it works effectively in providing explicit spatial clarity, this illustration can only communicate the ants' full emotional significance when paired with accompanying text: "At that moment, the two ants felt happier than they'd ever felt before. This was their home, this was their family. This was where they were meant to be". Once again, this text-image interplay creates a unified narrative, with each mode emphasizing the lesson the ants have learned and their renewed appreciation for the safety and connection of their community. With its shifting interplay between words and illustrations, *Two Bad Ants* exemplifies "multimodal narratives," where meaning emerges from the interaction of different semiotic modes (Nikolajeva & Scott, 2001). By combining textual and visual storytelling, this narrative technique invites older readers to explore subtle interactions between the two storytelling mediums while also remaining accessible to younger readers through visual clues.

Voices and Views: Text-Led Storytelling in *Dear Diary*
Dear Diary has been well received by readers for its playful and inventive design. The story is told in the form of diary entries and explores how the events in a single day are experienced from multiple perspectives. In this picture book, each of the eight narrators takes turn presenting the same sequence of events from their own point of view, offering a highly subjective and humorous account of

the day. Unlike *Two Bad Ants*, where words and illustrations share equal narrative responsibility, *Dear Diary* is clearly a text-led picture book, with illustrations playing a secondary yet still crucial role. This dynamic illustrates an enhancement interaction, where one of the two elements does not introduce new or contradictory information against the other, but instead works in strengthening the emotional or thematic depth of the story. To understand this interplay between words and illustrations, the role played by each component should be closely examined.

At first glance, the illustrations without accompanying text appear disjointed and almost incomprehensible due to their lack of sequential logic and cohesion. As the story unfolds, the focus of the illustrations on each page shifts from one character to another, beginning with a young girl, then a chair, a spider, a firefly, a knife and a fork, a dog, and finally a lady bird. Without the guiding presence of the diary entries, it becomes extremely difficult for readers to piece together the events and understand the relationships between different characters. These fragmented illustrations only come together to form cohesive diary entries when paired with contextual details supplied by text. In other words, it is the text that provides narrative coherence and continuity, guiding readers to understand that the story begins with the first diary of a young girl called Lucy telling about her day at school, taking her dog Bubu for a walk and her parents' party in the evening. This is followed by the second diary of a chair being overturned in a classroom and then another six more entries from other characters involved in the course of Lucy's day. In addition to providing thematic and chronological links between each illustration, it is also the text that helps readers recognize how each diary entry is presented with the characters' unique feelings, biases, and interpretations of the same events. For example, while Lucy sees her actions as playful, the chair narrates the same moment with a tone of frustration, highlighting its discomfort after being knocked over in the classroom. In this text-dominant picture book, the text thereby functions as the "narrative glue" (Nikolajeva & Scott, 2001), guiding readers in interpreting the visual elements and providing the structural coherence necessary to unify the disparate perspectives.

Although secondary to the text, which acts as the driving force of the story, the illustrations also work effectively in reinforcing the tone and emotional perspective of each narrator. Throughout the story, distinctive collage style featuring mixed media, varied textures, and expressive hand lettering is employed to reflect the emotions and personalities of each narrator. For example, the ladybird is depicted with wings bursting in vibrant colors and swirling patterns to reinforce her energetic and self-confident tone (See Figure.4). In contrast, the chair's diary entry is illustrated with rigid, angular shapes that evoke its sense of stiffness and discomfort, corresponding with its frustrated reaction to being overturned expressed in the text (See Figure.5).

Figure.4 Dear Diary
(Source: Stickers and Stuff, 2011)

Figure.5 Dear Diary
(Source: Stickers and Stuff, 2011)

In addition, different background materials and layouts are used in each diary entry to emphasize each narrator's individuality. For instance, the knife and fork's diary entry is set on a tidy grid background, with clean lines and orderly layout to reflect their formal, slightly uptight personalities. Their serious expressions and stiff body language match their anxious tone as the text describes about being separated for cleaning (See Figure.6). On the other hand, the dog's diary is presented on a messier background, with his image running across the double-page spread, surrounded by playful, childlike handwriting. This chaotic visual style effectively reflects the dog's excitement, confusion, and longing for attention (see Figure.7).
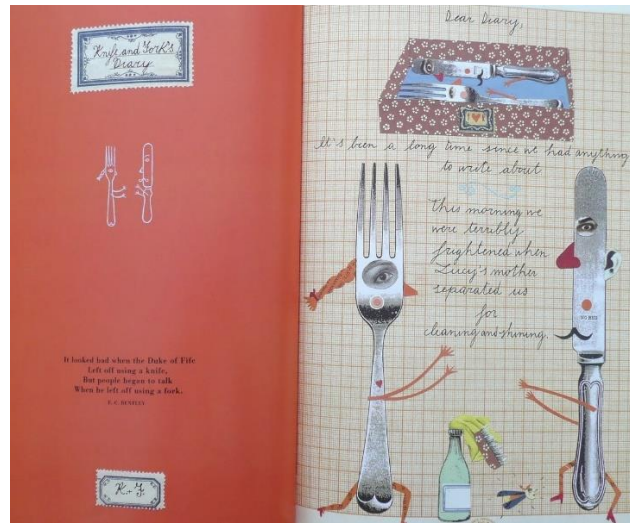
Figure.6 Dear Diary
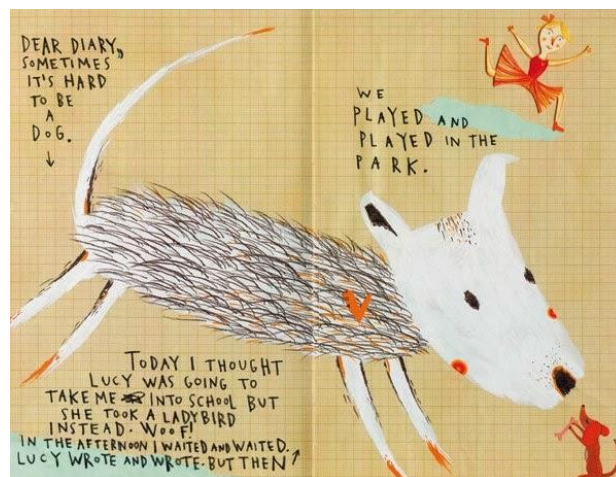(Source: Stickers and Stuff, 2011)

Figure.7 Dear Diary
(Source: Stickers and Stuff, 2011)

As discussed above, it is obvious that visual elements in *Dear Diary* do not introduce new narrative content or clarify plot details, but they heighten the mood, deepen character voice, and support the humor embedded in the text. This stylistic approach exemplifies enhancement interaction, in which the illustrations amplify the emotional or thematic qualities of the story without presenting independent or contradictory information. While the text provides structure, context, and narrative continuity, it is the illustrations that bring the characters' feelings and personalities to life in ways that words alone cannot fully express. Without the text, the story would lack coherence and sequence. Conversely, without the illustrations, the story would lose much of its emotional color and imaginative charm. The interaction between these two elements thereby invites readers to experience each diary entry on multiple levels, combining verbal and visual clues to create a richer, more engaging reading experience.

Imagery in Motion: Visual-Only Narratives in *Sector 7*

Compared to the two previous picture books, *Sector 7* demonstrates how image on its own can drive an entire narrative without additional supporting text. This wordless picture book takes readers into a creative and imaginative world in which visual elements become the focus of creating stories. The story follows a young boy on his school fieldtrip to the Empire State Building. There, he becomes friend with a cloud who takes him to the cloud factory in the sky, known as Sector 7. Upon arrival, he realizes the clouds are not happy with how they are shaped and patterned. With his creativity, the boy introduces new designs for the cloud. While adults at Sector 7 are not happy about this change, the clouds themselves are excited for the new possibilities. With no text, this picture book relies solely on visual clues to navigate its storyline, with the illustrations carrying the full load of storytelling. It is therefore interesting to consider how this meaning emerges through visual means, and how the reader becomes a co-creator of the narrative.

Unlike a picture book with text, it can be difficult to create a cohesive story in a wordless picture book since the visuals alone must clearly establish characters, settings, and plot. In *Sector 7*, this challenge is resolved through elaborate, high modality illustrations, ranging from the realistic depictions of human characters and urban landscapes to the portrayals of expressive clouds in the sky. Whereas in *Dear Diary* the text provides the anchor for the narrative, the illustrations in *Sector 7* themselves independently convey a straightforward story of events in chronological sequence. For example, directional clues such as a cloud carrying the boy left to right indicate forward progression, while the school bus that takes him home moves right to left, suggesting the end to a day's adventure. As Nikolajeva and Scott (2001) noted, this kind of visual sequencing within wordless picture books acts as the main device used to direct readers through the story in a smooth manner.

Additionally, subtle narrative markers embedded in the illustrations provide clues that help readers understand the story. For example, the boy's drawing of sea creatures on a frosted bus window in one of the opening scenes immediately conveys the boy's artistic talent, which plays an important foreshadowing role to the boy's creative contributions at Sector 7. Visual coherence continues because the boy then uses design tools at the cloud factory, composing fanciful cloudscapes based on the earlier motifs (See Figure.8). Moreover, the characters' facial expressions effectively express emotional nuance. For instance, the boy's initial disappointment when he finds the observation deck contrasts sharply with the clouds' joyful expressions when their new designs are revealed. These visual clues related emotion stimulate more engaged reading by enabling readers to interpret characters' internal emotions and states, as proposed by Arizpe and Styles (2016). Throughout the story, the readers are also visually placed in a vividly imaginative yet familiar setting. For instance, the observation deck of the Empire State Building has a signpost with the label "Observatory," placing readers in a familiar place. Likewise, the cloud factory itself appears as a train station, with arrivals and departure screens, a check-in station, and waiting areas. In doing so, this design not only establishes the purpose of the factory, but also conveys a sense of bureaucracy and order (See Figure.9).

Figure.8 Sector 7
(Source: Polita, 2022)

Figure.9 Sector 7
(Source: Polita, 2022)

Another interesting aspect of *Sector 7* is its ability to attract an audience that is referred to as a "dual audience" (Nikolajeva & Scott, 2001). As a wordless picture book, this story effectively communicates with readers on multiple levels. For example, younger readers may relish the straightforward visual storytelling and the details, such as the clouds in fanciful shapes and the playful exchanges between boy and clouds. At the same time, adult readers can enjoy deeper thematic elements, such as the conflict between creativity and bureaucratic rigidity that is visually expressed by the mechanization of the cloud factory. However, despite is strengths, a wordless picture book may pose certain interpretive challenges, especially if for younger readers who have not yet read complex visual narratives. With no textual instructions, readers must rely solely on visual literacy skills to grasp the complete narrative. For example, recognizing the boy's artistic talent in the opening scene and identifying Sector 7 as a cloud dispatch center requires careful attention to detail and contextual interference.

As Sipe (2012) points out, wordless picture books often encourage multiple readings, as readers discover new details and layers of meaning with each revisit. The strategic use of recurring visual motifs and interconnected illustrations in this story is therefore used in this picture book to ensure that each reading offers new discoveries.

From the above analysis, *Sector 7* makes it rather clear that illustrations alone can work independently to convey an entire story, unlike *Two Bad Ants* and *Dear Diary*, which rely on interplay between text and illustrations. This underlines the increasing significance of visual literacy, which is the capacity to comprehend visual information. Wordless picture books like *Sector 7* make it possible for students to practice observing details, thinking in creative ways and discussing ideas in an educational context. Students who develop visual literacy skills are better equipped to navigate a visually-rich world.

Conclusion

This study demonstrates how different narrative strategies can be employed in picture books through varying interactions between words and illustrations, which uniquely shape readers' engagement and interpretive possibilities. *Two Bad Ants* exemplifies how complementary and counterpoint interactions can work together in creating deeper meaning, allowing readers to imagine ordinary environments through the characters' perspectives. Conversely, *Dear Diary* shows how illustrations can effectively enhance a text-driven picture book by amplifying tone, personality, and humor, enriching readers' emotional experience without competing with the text. By contrast, *Sector 7* demonstrates how narrative complexity can be conveyed independently by illustrations, emphasizing readers' interpretive freedom and the growing necessity of visual literacy in today's visually-driven world.

Beyond literary analysis, these narrative strategies carry important implications for both educational practices and creative work. Understanding how the interplay between text and illustrations operates can help educators design instruction that supports students' verbal and visual literacy development in integrated and engaging ways. For example, teachers may select picture books like *Sector 7* to foster students' observation skills, inferential thinking, and critical visual analysis, while also using *Dear Diary* or *Two Bad Ants* to encourage exploration of emotional expression, narrative point of view, and visual–textual interplay. Picture books can also be incorporated into interdisciplinary classrooms, such as art, language, and media studies, to promote multimodal thinking and collaborative learning through visual storytelling. Additionally, guided reading activities and comparative analysis tasks using these books can cultivate students' interpretive skills and creative responses. For picture book creators, an awareness of the diverse forms of text–image interaction can enrich narrative techniques and stylistic choices, allowing for more purposeful design that addresses various developmental stages and learning needs.

Future research could further investigate how specific types of text–image relationships influence children's comprehension, engagement, and creativity across different age groups or cultural contexts. Longitudinal studies might also examine how sustained exposure to multimodal narratives contributes to broader literacy development and aesthetic appreciation.

References

Arizpe, E., & Styles, M. (2016). *Children reading pictures: Interpreting visual texts* (2nd ed.). Routledge.

Fanelli, S. (2000). *Dear diary*. Candlewick Press.

Goss, R. (2020, June 15). *Two bad ants by Chris Van Allsburg.* https://www.youtube.com/watch?v=f4H zn1G9410.

Lewis, D. (2001). *Reading contemporary picturebooks: Picturing text*. Routledge Falmer.

Lynch-Brown, C., & Tomlinson, C. M. (2002). *Essentials of children's literature* (4th ed.). Allyn & Bacon.

Nikolajeva, M., & Scott, C. (2001). *How picturebooks work*. Routledge.

Nodelman, P. (1988). *Words about pictures: The narrative art of children's picture books*. University of Georgia Press.

Pantaleo, S. (2008). *Exploring students' responses to contemporary picturebooks*. University of Toronto Press.

Polita, M. (2022, December 11). *2022 Christmas Gifts – Day 11 – Search the Sky*. https://www.scaffalebasso.it/ david-wiesner -sector-7-orecchio-acerbo/

Sipe, L. R. (2012). Revisiting the relationships between text and pictures. *Children's Literature in Education, 43*(1), 4–21.

Stephens, J. (1992). *Language and ideology in children's fiction*. Longman.

Stickers and Stuff. (2011, February, 19). *Dear Diary - Sara Fanelli*. https://stickersandstuff. blogspot.com/2011/ 02/dear-diary-sara-fanelli.html

Van Allsburg, C. (1988). *Two bad ants*. Houghton Mifflin Harcourt.

Wiesner, D. (1999). *Sector 7*. Clarion Books.