Reading Comprehension in AI-Supported Storytelling: Evidence from Fourth-Grade Students' Interactions with Generative Artificial Intelligence

Mehmet Emrah Kuru*
Canakkale Onsekiz Mart University (Türkiye)

Mustafa Yunus Eryaman**
Canakkale Onsekiz Mart University (Türkiye)

Abstract

This study investigates how fourth-grade primary school students' interactions with generative artificial intelligence tools—specifically ChatGPT and Microsoft Copilot—during story-creation tasks are reflected in their native-language reading skills. Using a qualitative design embedded within an action-research process, data were gathered through semi-structured interviews, in-class observations, field notes and students' written responses to two story-comprehension forms derived from AI-generated narratives. The data were analysed through thematic content analysis, supported by descriptive statistics. Guided by Rosenblatt's reader-response theory, Kintsch's construction–integration model, inference-based comprehension research and the Turkish fourth-grade language arts learning outcomes, the analysis produced five key themes: Text Comprehension, Inference-Making, Detail Awareness, Emotional Response and Connection, and Word Consciousness. Across the dataset, 132 codes were identified. Findings indicate that students demonstrate strength in fundamental text comprehension and event-focused interpretation but show limited development in abstract inference, vocabulary awareness and deeper affective engagement. AI-supported story production appears to scaffold comprehension processes by helping students recognise narrative structure, form empathetic connections with characters and draw on personal experiences while making meaning. Individual analyses of focal participants reveal a shift from surface-level recall to more reflective, value-oriented and empathy-driven reading behaviours. Overall, the results suggest that AI-supported storytelling holds promise as a pedagogical tool for enhancing native-language reading, provided it is accompanied by explicit instruction targeting inference-making, vocabulary development and critical engagement with AI-generated texts.

Keywords: Artificial Intelligence, Generative AI, Reading Comprehension, Primary Education, Turkish Language Arts

* *Mehmet Emrah Kuru*, Canakkale Onsekiz Mart University, School of Graduate Studies, Department of Elementary Education, ORCID ID: 0000-0002-8836-8349

** *Prof. Dr. Mustafa Yunus Eryaman*, Canakkale Onsekiz Mart University, Faculty of Education, Department of Elementary Education, ORCID ID: 0000-0002-8836-8349

Correspondence: *m.emrahkuru@hotmail.com*

Introduction

The rapid development of generative artificial intelligence (AI) has introduced new and still largely unexamined dynamics into language and literacy education. Tools such as ChatGPT and Microsoft Copilot can instantly generate coherent narrative texts from brief prompts, making it possible to design reading tasks that are seemingly authentic, engaging, and closely aligned with students' own narrative contributions. Despite this potential, the specific consequences of these tools for children's native-language literacy, especially in the primary school years when foundational reading skills and dispositions are formed, remain insufficiently understood. This lack of clarity becomes more critical when AI is integrated directly into core processes of reading and comprehension. When students not only read but also co-construct texts with AI systems, the nature of their interaction with written language changes: reading evolves from a one-way encounter with a pre-given text into a technology-mediated, iterative process. In such contexts, it is not yet known how AI-mediated narratives shape students' comprehension of stories, their inferential thinking, their attention to textual details, their emotional engagement with characters, or their sensitivity to words and language use.

Existing research on AI in education suggests that these technologies have the capacity to transform reading practices at a structural level. Ademola's (2024) comprehensive study shows that AI-supported learning environments fundamentally reorganise traditional reading strategies and the pedagogical assumptions that underpin them. According to this study, AI integration allows instructional processes to be personalised at a micro level: intelligent systems can monitor students' reading performance in real time, detect emerging difficulties in textual comprehension, and provide immediate, data-driven feedback. Such capabilities create new possibilities for targeted scaffolding and for preventing the consolidation of misunderstandings.

The integration of AI into reading processes offers powerful mechanisms for tailoring learning to individual differences. Adaptive learning systems can adjust content and task demands to students' cognitive profiles and learning preferences, thereby deepening the quality of interaction with texts. By analysing learners' comprehension levels dynamically and reorganising reading activities on the basis of these analyses, AI algorithms can increase the apparent effectiveness and precision of educational interventions. Yet, these opportunities are intertwined with substantial ethical, epistemological, and professional challenges. As Ademola emphasises, issues related to data privacy, unequal digital access, and algorithmic bias in AI systems' data collection and analysis processes pose serious risks. If left unaddressed, such risks may reinforce existing inequities and undermine trust in educational technologies. Moreover, the expanding role of AI redefines teachers' professional responsibilities: it is no longer sufficient for educators merely to operate AI tools; they must also develop the pedagogical and critical competencies required to integrate these tools in ways that genuinely enhance literacy learning rather than displace it. This, in turn, points to the need for teacher education programmes that include AI literacy and technology-enhanced pedagogical approaches.

Taken together, these considerations highlight a central tension: while AI technologies, and generative AI in particular, offer unprecedented opportunities to personalise and enrich reading experiences, their concrete impact on students' native-language reading comprehension—especially in authentic classroom contexts—remains an open empirical question. Addressing this question is crucial for ensuring that AI-supported reading innovations are pedagogically meaningful and ethically sustainable. Against this backdrop, the present study investigates how fourth-grade students' interactions with AI-based storytelling tools are reflected in their reading comprehension in Turkish. Within a structured classroom implementation, students provided narrative elements (character, setting, time and plot) to ChatGPT and Microsoft Copilot, obtained AI-generated stories based on these inputs, read the resulting texts, and then completed story-comprehension forms. The analysis focuses on how these AI-mediated experiences relate to students' abilities to understand narrative texts, make inferences, attend to details, develop emotional connections with characters, and demonstrate word consciousness. By examining these dimensions, the study aims to contribute empirical evidence to ongoing debates about the role of generative AI in reading education and to

inform future pedagogical and policy decisions concerning AI-supported literacy practices. The study is guided by the following research question:

How are fourth-grade primary school students' native-language reading skills reflected in their responses to stories generated through interaction with AI tools such as ChatGPT and Microsoft Copilot?

Theoretical Framework

The analytical framework for this study is grounded in three complementary theoretical perspectives and the national curriculum, each illuminating a distinct dimension of students' reading processes in AI-mediated learning environments. Rosenblatt's (1938, 1978) reader-response theory provides the affective and experiential foundation, emphasising the transactional nature of reading and the ways in which meaning emerges through the dynamic interplay between the reader and the text. From this perspective, students' prior experiences, emotional states and personal associations are not peripheral but essential elements that shape their engagement with AI-generated narratives—texts that are simultaneously familiar and novel due to their algorithmic authorship.

Kintsch's (1998) construction–integration model contributes a cognitive lens, foregrounding the processes through which readers build propositional representations, integrate local and global textual cues and construct coherent situation models. This framework is particularly relevant for understanding how children process AI-generated stories, as such texts often display structurally consistent narrative patterns that may either support or challenge young readers' abilities to detect causal relations, track narrative progression and maintain global coherence.

Inference-based models of reading comprehension (e.g., Cain & Oakhill, 1999) inform the study's focus on higher-order meaning-making. These models highlight the centrality of bridging and elaborative inferences for moving beyond literal interpretations—an especially salient competency when interacting with AI-produced texts that may contain implicit cues, culturally nuanced expressions or atypical narrative decisions. This perspective guided the identification of students' attempts to interpret unstated meanings, derive value-based messages and engage critically with narrative content.

In addition to these theoretical foundations, the study is aligned with the Turkish Ministry of National Education (MEB, 2024) fourth-grade Turkish language arts outcomes, especially those pertaining to:

- identifying characters, events, setting, main idea and theme

- making predictions and inferences from text

- analysing emotional and value-laden messages

- recognising and using words with varied semantic relations

Together, these curricular and theoretical perspectives informed the development of the coding framework and ensured that the thematic categories reflected not only data-driven patterns but also established models of reading comprehension, cognitive processing and reader–text interaction within AI-supported literacy contexts.

Methodology

Research Design

The study employs a qualitative approach embedded in an action-research cycle. Conducted over a period of 45 days, the process was iterative: initial AI-supported story tasks were implemented,

student responses were analysed, and subsequent tasks—including the second, more structured comprehension form—were refined to elicit deeper and more reflective reading behaviours.

Study Group and Context

The participants were fourth-grade students who took part in the action-research process conducted in a public primary school. All participating students engaged in AI-supported story-creation tasks throughout the implementation. A total of 30 students participated in the study, and six focal participants (Ezgi, Fatma, Bengü, Önder, Umut and Ahmet) were selected for in-depth qualitative analysis. All participant names are pseudonyms assigned to protect student identity. These six key participants were chosen through maximum variation sampling, taking into account differences in participation, expressive behaviours, socio-economic background, family structure and access to technology.

For the purposes of in-depth qualitative analysis, these focal students were examined closely because they represented a broad range of engagement styles, narrative preferences and developmental characteristics relevant to the aims of the study. The key participants attended a fourth-grade classroom in a public primary school located in a city of Marmara Region, Türkiye. All were approximately ten years old and came from diverse socio-economic backgrounds, family structures and levels of access to digital resources. This diversity enriched the AI-supported storytelling process and enabled a wide range of reading behaviours to emerge.

Fatma lives with her parents and two older sisters; her father works in animal husbandry and farming, while her mother occasionally supports these activities. The family has a laptop and internet connection at home, although Fatma does not have her own tablet or phone. Both parents are middle-school graduates. Fatma enjoys drawing, singing, dancing and caring for cats.

Bengü's mother is a homemaker and her father is a farmer. She has one older sister and one older brother. Although the household does not have a computer, they have internet access. She has a smartphone without an active SIM card. Her mother completed primary school and her father completed middle school. She enjoys drawing, singing and dancing, and her family keeps a cat, dog and several birds at home.

Ahmet's parents are farmers; his mother completed primary school and his father did not finish middle school. The household has internet access and a tablet but no computer or phone. Ahmet is an only child and enjoys singing, playing ball games and hide-and-seek. He reported an accident in his early childhood that he still remembers vividly, which appears to contribute to his strong attention to detail in narratives.

Ezgi's father works in the industrial sector and her mother is a homemaker. She has one older sister and one younger brother. The family has a computer and internet access at home, and Ezgi also owns a tablet. Her mother is a high-school graduate and her father is a university graduate. She enjoys dancing, playing volleyball and singing, and she keeps pigeons at home, showing a strong interest in animals.

Önder's father works in the industrial sector and his mother takes care of cattle. He has one older brother. Although the family does not have a computer, they have internet access, and Önder owns a tablet. He keeps pigeons in the garden and enjoys playing ball games and spending time outdoors.

Umut's father works in the industrial sector and his mother is a homemaker; both parents are middle-school graduates. He has one older sister and one younger brother. The family has a computer and internet connection; Umut has limited internet access on his smartphone but does not own a tablet. He enjoys drawing, and the family keeps a cat. Umut mentioned experiencing a boat accident in the past, which appears to have heightened his emotional sensitivity in narrative tasks.

None of the students had prior experience with AI tools before the study. The intervention therefore constituted their first exposure to AI-supported storytelling, creating a naturalistic learning environment in which novice users engaged with generative AI in authentic classroom conditions.

Data Collection

To ensure methodological rigor and establish robust data triangulation, the study utilised multiple complementary data sources.

AI-supported story-creation tasks:

Students generated narratives using two generative AI tools. First, they interacted with ChatGPT (GPT-3.5), followed by Microsoft Copilot, by providing narrative elements such as character, setting, time and plot. These tasks created a naturalistic environment in which students engaged with AI-mediated storytelling for the first time, producing texts that served both as instructional materials and as data for subsequent comprehension analyses.

Story-comprehension forms (Form 1 and Form 2):

Form 1 was adapted from Esmer's (2019) validated narrative text comprehension test for primary students and was revised through expert consultation to align with the aims of the present study. It assessed foundational narrative components (characters, time, place, events and problem–solution structures). Form 2, developed specifically for this research, included more elaborative and cognitively demanding prompts targeting inference-making, emotional engagement and word consciousness. Its design was informed by the patterns emerging from Form 1 and by expert feedback, enabling the elicitation of deeper, more reflective reading behaviours.

Observations and field notes

The researcher conducted systematic classroom observations throughout both the AI-interaction and reading-comprehension sessions. Student behaviours, engagement levels, spontaneous verbal expressions and peer interactions were documented in detailed field notes to capture contextual and process-oriented dimensions of reading.

Semi-structured interviews:

Interviews were carried out to further explore students' interpretations of AI-generated stories, their perceptions of characters and events and their experiences using AI tools. These interviews enriched the data set by offering insight into students' cognitive, emotional and evaluative reasoning.

Data Analysis

All qualitative data sources—including students' story-comprehension forms, semi-structured interviews, field notes and supplementary documents—were analysed through a thematic content analysis procedure. The analytic process followed a two-stage design that integrated inductive and framework-based strategies, ensuring that emerging themes were simultaneously grounded in the empirical data and theoretically coherent with the study's conceptual foundations.

In the first stage, an inductive, data-driven coding process was undertaken. Meaningful units were identified directly from students' written and verbal responses, as well as their observable reading behaviours, without the imposition of predetermined categories. This stage enabled nuanced patterns to surface organically—for instance, how students interpreted narrative events, formed inferences, expressed emotional reactions or attended to specific textual cues. Rather than producing a simple descriptive inventory, this inductive phase constituted a rigorous bottom-up content analysis in which initial codes were derived from students' authentic expressions and behavioural indicators. This

approach ensured that subsequent analytical categories faithfully reflected the lived reading processes mediated by AI-generated narratives.

Second, the analysis proceeded with framework-based coding. The inductively generated codes were reorganised within an analytical structure derived from the theoretical framework and the MEB (2024) Turkish reading outcomes. This framework drew specifically on:

• Rosenblatt's (1978) reader-response theory (emotional engagement, aesthetic stance),

• Kintsch's (1998) construction–integration model (global and local text comprehension),

• Cain and Oakhill's study on inference-making processes,

• international research on word consciousness and metalinguistic awareness,

• MEB (2024) reading outcomes relating to comprehension, inference, attention to details and vocabulary development.

Through this two-stage analytic process, the resulting themes were not merely descriptive categories; rather, they reflected higher-level conceptual patterns that explained students' cognitive and affective reading processes. The themes emerged inductively from the data and were subsequently strengthened and clarified through alignment with theoretical models.

Identified Themes

1. *Text Comprehension*: Students' ability to construct both global and local meaning from narrative texts, consistent with Kintsch's macro- and microstructure processes and MEB comprehension outcomes.

2. *Inference-Making*: Students' use of bridging and elaborative inferences, aligned with the cognitive mechanisms described in Cain and Oakhill's inference research.

3. *Detail Awareness:* Sensitivity to descriptive elements, narrative cues and contextual details necessary for coherent story understanding.

4. *Emotional Response and Connection*: Affective engagement with characters and events, reflecting Rosenblatt's conceptualisation of aesthetic reader–text interaction.

5. *Word Consciousness*: Awareness of vocabulary, word meanings and language use, consistent with research on word consciousness and metalinguistic development.

Trustworthiness of the Analysis

To enhance trustworthiness, three independent experts reviewed the coding scheme and coded a selected subset of the data. Their feedback was used to refine theme definitions, check the alignment between inductive codes and the theoretical framework, and resolve any discrepancies. These procedures strengthened the credibility, dependability and confirmability of the analysis in line with qualitative research standards (Lincoln & Guba, 1985).

Findings

Table 1. *Frequencies and Percentages of Codes by Theme and Sub-Category*

Themes and Subcategories	F	Overall Percentage (%)	Percentage Within Theme (%)
Text Comprehension	52	39.39	100.0
→ Character	18	13.64	34.62
→ Time	10	7.58	19.23
→ Setting	6	4.55	11.54
→ Event	12	9.09	23.08
→ Main Idea	6	4.55	11.54
Inference-Making	28	21.21	100.0
→ Value Identification	15	11.36	53.57
→ Extraction of Main Message	13	9.85	46.43
Detail Awareness	24	18.18	100.0
→ Noteworthy Scene	14	10.61	58.33
→ Sequence of Events	10	7.58	41.67
Emotional Response and Connection	20	15.15	100.0
→ Character Preference	11	8.33	55.0
→ Empathy / Identification	9	6.82	45.0
Word Consciousness	8	6.06	100.0
→ New Word / Expression	8	6.06	100.0

Note. The frequencies presented in this table were determined through content analysis of data obtained from interview transcripts, observations, field notes and document analysis. The percentages were calculated based on both the overall frequency distribution and the proportional distribution within each theme.

Table 1 shows how the 132 codes generated through content analysis were distributed across the five themes as follows: Text Comprehension (39.39%), Inference-Making (21.21%), Detail Awareness (18.18%), Emotional Response and Connection (15.15%) and Word Consciousness (6.06%). This distribution indicates that students most frequently displayed behaviours related to basic text comprehension and narrative structure, whereas vocabulary-related awareness was relatively rare.

Within Text Comprehension, the sub-categories of Character (34.62% of this theme) and Event (23.08%) were more prominent than Time, Setting and Main Idea. This suggests that students tended to focus on "who" and "what happened" rather than "when," "where" or the abstract main idea.

The Inference-Making theme (21.21%) was split between Value Identification (53.57% within the theme) and Extraction of Main Message (46.43%), showing that students could produce value-laden comments (e.g., the importance of friendship) but still needed support for consistently uncovering implicit messages.

In Detail Awareness (18.18%), most codes related to Noteworthy Scene (58.33%), such as specific moments of conflict or turning points, while Sequence of Events (41.67%) reflected students' attempts to organise events chronologically.

Emotional Response and Connection (15.15%) captured students' preferences for certain characters (Character Preference, 55%) and their attempts at Empathy/Identification (45%).

Finally, Word Consciousness (6.06%)—entirely composed of New Word/Expression codes— was the least represented theme, suggesting limited spontaneous attention to vocabulary.

Text Comprehension: From Identifying Narrative Elements to Grasping Themes

Across the data set, students demonstrated a relatively strong ability to recognise fundamental narrative elements. For example, Ezgi correctly identified the characters ("Simay and Eslem"), time ("Thursday morning") and main activity ("playing with snowballs"), and labelled the main idea as "friendship." Similarly, other students were able to name characters and describe the setting as a park, a school yard or a ship, and to provide a basic description of what happened.

However, the analysis also showed that Time, Setting and especially Main Idea were less frequently and less elaborately addressed compared to character and event information. In many responses, the time and place were either omitted or expressed in very general terms, and the main idea was often reduced to a single word (e.g., "friendship") without explicit links to specific events in the story.

At the individual level, there was evidence of progression across the two forms. For instance, Ezgi's second-form responses moved beyond naming narrative elements to emphasising the overall theme of "close friendship and snow," indicating a shift from listing details to thematically organised comprehension.

Figure 1. Visual representation of Ezgi's responses to narrative text comprehension forms based on the ChatGPT-generated story

Figure 1 illustrates Ezgi's responses to the narrative text comprehension forms generated from the ChatGPT-produced story. Her answers show a clear understanding of the fundamental narrative elements, including characters, events, time and setting. Ezgi's progression from surface-level recall in the first form toward more theme-oriented and value-based interpretations in the second form is evident in the visual representation. She identified meaningful scenes, articulated personal connections with the characters and demonstrated emerging skills in inference-making by linking story events to broader ideas such as friendship and shared experiences. Although her word consciousness remained limited, Ezgi's overall engagement reflects a developing ability to construct deeper meaning from AI-generated narratives through both cognitive and emotional processes.

Inference-Making: Emerging but Uneven Higher-Order Understanding

The theme of Inference-Making captured students' attempts to derive meanings that were not explicitly stated. Many students were able to express general values such as "spending time with friends is precious" or "teachers should be treated kindly," indicating that they were beginning to interpret the moral or social messages of the stories.

For example, Ezgi's statement that "the time I spend with my friends is very valuable" shows a value-based inference that connects the narrative to her own social experiences. Similarly, Fatma's shift from a general statement ("We should be good to teachers") to a more story-bound theme ("playing together") suggests increased alignment between the narrative events and the extracted message.

Figure 2. Visual representation of Fatma's responses to the narrative text comprehension forms based on the ChatGPT output
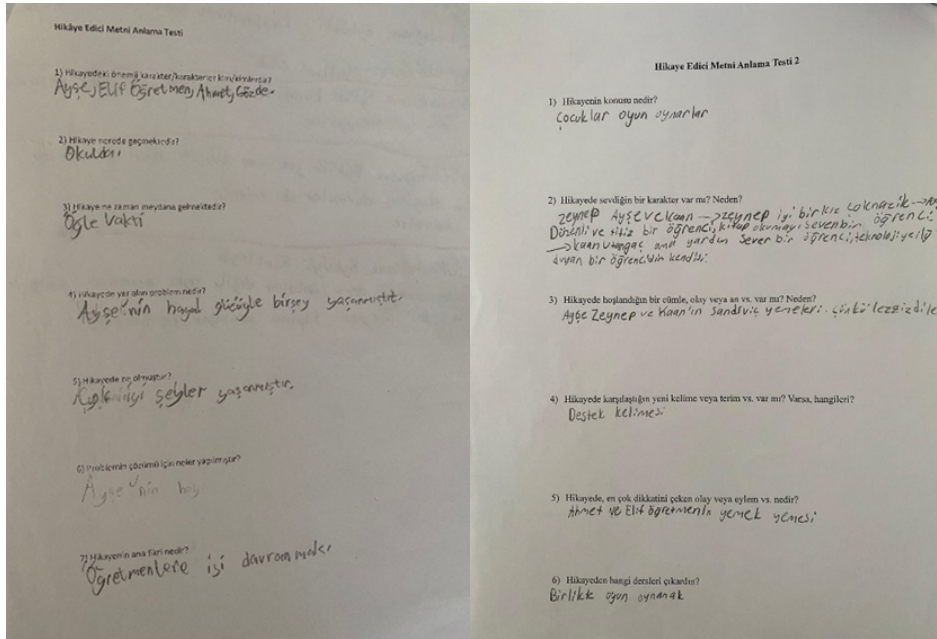
Figure 2 presents a visual summary of Fatma's responses to the narrative text comprehension forms based on the ChatGPT-generated story. The figure shows that Fatma successfully identified key narrative elements and highlighted specific scenes that captured her attention. Her responses reflect an emerging ability to infer simple messages from the story and to justify her character preferences with value-based reasoning. However, the visualisation also indicates limited consciousness of new vocabulary, suggesting that her lexical development remained relatively constrained during the process.

Nonetheless, across the sample, inferential responses were often brief, sometimes formulaic and not always anchored in specific textual evidence. Students could name abstract themes (e.g., friendship, cooperation, overcoming difficulties) but rarely elaborated on how particular scenes or character behaviours supported those themes. This indicates a need for structured support in linking textual details to abstract interpretations.

Detail Awareness: Attention to Scenes, Turning Points and Atmosphere

The Detail Awareness theme revealed that students were able to notice and highlight salient scenes, especially those involving conflict, resolution or vivid imagery. Their responses frequently referenced concrete moments such as:

- "Simay breaking the snow wall"
- "Meeting at the park together"
- "Ahmet and Elif Teacher eating together"
- "Erva hiding under the bridge"
- "The ship surviving the storm without damage"

These examples show that AI-generated stories, with their rich event-driven content, supported students in identifying narrative turning points and key scenes. In several cases, students not only recalled such details but used them to reconstruct the atmosphere of the story, for instance by referring to laughter during snowball fights or shared activities such as football and picnics.

However, students' explanations of event sequence were more variable. Some could correctly order events and indicate cause–effect relations, while others provided partial or fragmented sequences. The second, more structured comprehension form appeared to help students articulate

event sequences more clearly, suggesting that task design plays a crucial role in eliciting detailed narrative awareness.

Table 2. *Thematic Content Analysis of Fatma's Narrative Text Comprehension Process*

Theme	Sub-Codes	Example Student Expression
Text Comprehension	Time, setting, event, character	"While Teacher Ahmet was sitting on the bench…", "They played games", "Ayça's closest friends…"
Inference-Making	Extracting the main idea, generating meaning	"Being kind to teachers." — "Playing together."
Detail Awareness	Striking scene, event, or expression	"Ahmet and Teacher Elif eating together." — "Ayşe, Zeynep and Kaan eating sandwiches."
Emotional Response and Connection	Preferred character and justification	"Zeynep is a good girl, very polite." — "Kaan is shy but helpful."
Word Consciousness	Newly encountered word	"The word 'destek' (support)."

Table 2 summarises the thematic content analysis of Fatma's responses during the narrative text comprehension process. The data show that she successfully identified key narrative elements such as time, setting, events and characters, demonstrating competence in basic text comprehension. Her inferential statements, though brief, indicate an emerging ability to derive main ideas and generate meaning from events. Fatma also displayed attention to salient details, referencing specific scenes that captured her interest. In terms of emotional engagement, she justified her character preferences with value-based reasoning, reflecting her developing capacity for empathic connection. However, her consciousness of new vocabulary remained limited, suggesting that word consciousness continues to be an area requiring instructional support.

Emotional Response and Connection: From Listing Characters to Empathic Engagement

One of the most notable contributions of AI-supported storytelling to reading experiences was in the area of Emotional Response and Connection. As students engaged with characters who resembled their peers and everyday situations, they gradually moved from merely listing names to evaluating and identifying with characters.

For instance, Ezgi referred to a long-term friendship ("She has been my friend since childhood, so I love her very much"), signalling a strong personal connection between the story and her own life. Fatma described Zeynep as "a nice girl, very polite" and Kaan as "shy but helpful," showing that she could infer character traits and express positive evaluations. Bengü emphasised that she liked all the characters because "they behave well," highlighting a global value judgement about social behaviour.

In some cases, students explicitly expressed identification with a character, as in Önder's statement "Ömer is me," which indicates a high level of projection and empathy. Umut chose Emre because he prepared first-aid materials before the trip, appreciating the character's responsibility and prudence.

At the same time, some responses revealed ambivalence or critical distance. For example, Ahmet both selected a character he liked ("Önder, because he found those who were hiding") and stated that he did not like the roles of the characters overall, suggesting the emergence of a more evaluative and critical stance towards the narrative.

Overall, the findings show that AI-generated stories facilitated affective engagement and helped students relate textual events to their own experiences, especially in the second comprehension form where questions explicitly prompted personal reflection and justification.

Table 3. *Thematic Content Analysis of Önder's Narrative Text Comprehension Process*

Theme	Sub-Codes	Example Student Response
Text Comprehension	Character, time, event, main idea	"All of them"; "night"; "love of friendship"; "friendship"
Inference-Making	Value identification, sensing the main message	"Friendship is sacred"; "the discovery of the lost island"
Detail Awareness	Noteworthy event, sequence of events	"the discovery of the lost island"; "Ömer finding it"
Emotional Response and Connection	Empathy, identification with characters	"All of them, for no reason"; "Ömer is me"
Word Consciousness	New word, word awareness	"None"

Table 3 summarises the thematic patterns observed in Önder's narrative text comprehension process. His responses demonstrate strong recognition of basic story elements and a clear ability to identify value-based messages, particularly around themes of friendship. The data also show that he attended to key events and displayed notable emotional engagement by identifying himself with a character ("Ömer is me"). However, his responses indicate minimal attention to vocabulary, suggesting limited development in word consciousness compared to the other dimensions.

Word Consciousness: A Weak but Developable Dimension

Among the identified themes, word consciousness emerged as the least prominent dimension. A considerable number of students explicitly stated that the stories contained *"no new words,"* which suggests two possibilities: either the linguistic complexity of the AI-generated narratives largely matched their existing vocabulary repertoire, or students did not naturally orient their attention toward lexical novelty during the reading process.

Nevertheless, several noteworthy exceptions indicate that the potential for vocabulary enrichment was present. Fatma, for instance, identified *"destek"* (support) as an unfamiliar or salient term, while Umut highlighted *"egzotik"* and *"biyolojik"* as newly encountered words (see Figure 3). These isolated instances demonstrate that AI-generated texts did occasionally introduce new lexical items and that some students were capable of recognising and articulating these items when appropriately prompted.

The overall scarcity of such responses underscores the importance of structured instructional scaffolding. The findings suggest that vocabulary development within AI-supported reading environments may require more intentional pedagogical integration—for example, guiding students to mark unknown words, infer possible meanings through contextual cues, discuss word usage collectively, and incorporate newly learned vocabulary into their own written productions. Such

strategies could enhance students' metalinguistic awareness and promote deeper engagement with the lexical features of AI-generated narratives.

Figure 3. Visual representation of Umut's responses to narrative text comprehension form 2, generated based on the ChatGPT-assisted analysis

Figure 3 illustrates how Umut engaged with key dimensions of narrative understanding, including literal comprehension, inference-making, detail awareness, and emotional connection. The visual layout highlights both the strengths and the developing aspects of Umut's reading processes, providing a concise overview of how AI-assisted prompts supported deeper engagement with the text.

Figure 4. Visual representation of Bengü's responses to the narrative text comprehension forms based on the ChatGPT output

Figure 4 illustrates Bengü's responses across both narrative text comprehension forms, showing her progression from basic recall of characters, time and events to more elaborative

interpretations. In Form 2, Bengü demonstrated clearer reasoning, stronger emotional evaluations of characters and increased attention to scene-specific details, particularly those involving conflict and action. This visual comparison reflects her shift toward more purposeful and reflective comprehension when engaging with AI-generated stories.

Table 4. *Findings on Bengü's Reading Skills Assessment*

Skill Area	Form 1	Form 2
Text Comprehension	She clearly stated characters (Simay, Elif Hüsna, Fatma Nur, Erva), setting (park), time (afternoon), event sequence and main idea.	She expressed the theme as 'valuable time together' and explained her favourite characters with reasons.
Inference-Making	She made a meaningful inference such as 'park and tree,' sensing the main message.	Her interpretation 'friendship is valuable' shows understanding of emotional and social meaning.
Emotional Response and Connection	She adopted the characters and evaluated fun and sharing positively.	She said she liked all characters because 'they behaved well.'
Detail Awareness	She stated the problem generally, but the solution was missing.	She wrote a striking detail such as 'Erva hiding under the bridge,' showing focus on details.
Word Consciousness	No direct expression of word awareness in this form.	She stated 'I have no new words,' showing awareness.

Table 4 summarises Bengü's performance across the five key dimensions of reading comprehension. The findings indicate that she demonstrated strong literal understanding of characters, time and events, while also showing the ability to generate value-based interpretations in the inference-making dimension. Her attention to salient narrative details—particularly the discovery scene—suggests developing sensitivity to plot structure. Moreover, Bengü displayed a high level of personal connection and empathy by identifying herself with a character, highlighting the affective engagement fostered by AI-generated stories. However, her responses show limited evidence of word consciousness, indicating that vocabulary-focused support may still be needed.

Discussion and Conclusion

One of the aims of this study was to examine the effects of AI-supported storytelling practices (ChatGPT, Copilot) on fourth-grade students' reading skills. The findings indicate that students showed notable progress in text comprehension, inference-making, detail awareness, and emotional connection, while improvement in the dimension of word consciousness remained limited. This section discusses the results in comparison with the literature on fluent reading and comprehension, highlighting similarities, differences, and implications for instructional practice.

Aşıkcan and Saban's (2021) ten-week action research study demonstrated that instructional practices such as repeated reading, echo reading, choral reading, readers' theatre and poetry- or song-based activities not only reduced reading errors but also enhanced reading speed, accuracy and prosodic competence. These findings align with Schwanenflugel et al.'s (2004) assertion that prosody serves as a mediator between word recognition and comprehension, and with Whalley and Hansen's (2006) findings that prosodic sensitivity is a strong predictor of reading development. Although fluency was not directly measured in the present study, students' progression from surface-level retelling to higher-order thematic and message-based interpretations in the second measurement, as

well as their improved ability to select scene-specific details, suggests that fluency components—particularly multiple readings and the use of model texts supporting prosody—play a role in deepening comprehension.

In this context, structurally coherent AI-generated texts can function as model texts that digitally approximate the prosodic cues offered in multimodal practices such as those used by Aşıkcan and Saban. The rhythmic, stress-based and syntactic markers embedded in such model texts appear to facilitate younger students' movement toward higher-order meaning units. Henkel's (2025) study in Ghana supports this observation, revealing that automatic speech recognition systems and large language models yield comprehension and oral fluency assessments highly consistent with human raters. This highlights the potential of AI-based assessment tools as reliable alternatives in resource-limited educational settings.

Furthermore, the findings on repeated reading effects are consistent with those reported by Pereyra (2025) and Yılmaz (2009), who observed gains in speed, accuracy, and comprehension. Samuels' (1997) theoretical framework also supports these outcomes, proposing that repeated reading automatises word recognition, thereby allowing learners to allocate cognitive resources to higher-level meaning-making processes. Likewise, Dessemontet et al. (2024) reported that repeated reading interventions indirectly enhance comprehension performance. In a similar vein, the experimental study by Elmaadaway et al. (2025) involving Alexa-based AI applications showed significant improvements in students' oral reading fluency and comprehension, and highlighted the motivational benefits of AI tools within the learning process.

Although students in the present study demonstrated clear improvements in higher-order comprehension indicators—such as theme/message extraction and selection of relevant details—the limited development observed in word consciousness is noteworthy. As Pereyra (2025) emphasises, pre-teaching vocabulary plays a crucial role in supporting comprehension. Therefore, while AI-supported storytelling increases engagement and depth of understanding, vocabulary growth may remain limited when targeted word instruction and repeated reading practices are not integrated into the process. Short, cyclic repeated-reading sessions built around AI-generated texts—incorporating semantic relations (synonyms, antonyms, near-synonyms), contextual clues, and example-sentence construction—are thus recommended.

In the action study conducted by Akyol and Çoban Sural (2020) with a student with special educational needs, reductions in reading errors, increases in speed, and improvements in comprehension indicated that technical interventions (word study, repeated reading, syllabication) yield stronger outcomes when combined with motivational factors. Similarly, findings from Borella et al. (2010) and Fielding-Barnsley (2000) emphasise that individual differences in attention, motivation, and strategy use are decisive in reading performance. In the present study, students' predominantly reader-centered responses and active participation during peer feedback and oral reading sessions align with the social-cognitive processes outlined in Guthrie and Wigfield's (2000) motivation model. However, motivational gains that are not accompanied by text-centered strategies (e.g., character mapping, plot-structure diagrams, evidence-comment matching) may lead to limitations in analytical reading outcomes.

Cruz Brotons' (2025) study on a GPT-based teacher support tool reinforces this conclusion, showing that tasks tailored to learners' individual needs and activities integrating emotional awareness provide more inclusive and motivating learning environments. Many studies in the literature similarly emphasise the reciprocal interaction between motivation and achievement (Baker & Wigfield, 1999; Morgan & Fuchs, 2007; Wigfield & Guthrie, 1997). In the present study, students' descriptions of AI-generated texts as "longer, more fluent, coherent and thoughtful," and their statements that they "felt inside the story," suggest that their awareness of textual structure increased and that they approached revision processes more consciously. This indicates a cyclical progression: increased motivation enhances text quality; improved text quality enhances comprehension; and deeper comprehension reinforces motivation. When considered alongside prosodically rich activities (poetry, song, readers'

theatre), this pattern aligns with that described by Aşıkcan and Saban (2021). Additionally, Chango García and Lugmania Borja's (2025) findings show that ChatGPT-based instructional strategies significantly strengthen inferential comprehension, further supporting results from the present study.

In conclusion, students showed substantial progress in theme/message extraction, detail selection, and emotional engagement; however, the limited improvement in word consciousness indicates that deficiencies emerge when vocabulary instruction and repeated reading are not jointly planned. While these contrasts are supported in the literature on meaning-making, fluent reading, and motivational processes, they also indicate that fourth-grade learning outcomes in Turkish require more specialised instructional conditions at the word-level domain.

Overall, AI is becoming embedded in reading instruction not as a temporary novelty but as a permanent component that transforms cognitive and affective processes. This transformation becomes meaningful when it does not diminish but elevates human critical thinking and ethical sensitivity. When teachers connect students with the depth of algorithms and texts, the learning process can move beyond mere transmission of information to a richer environment where the multiple voices of creativity become visible.

Recommendations

Findings from this study suggest that AI-supported storytelling can effectively strengthen students' reading comprehension—particularly in understanding narrative elements, making inferences, recognising key details and forming emotional connections—while vocabulary awareness and other micro-linguistic components show more limited development. Based on these results, several recommendations for instructional practice can be made.

First, instructional tasks should be structured to transform students' aesthetic engagement with AI-generated stories into evidence-based reasoning. Activities such as retelling the story, completing missing paragraphs, proposing alternative endings, adding scenes, or justifying interpretations with textual evidence can help students move beyond surface-level appreciation toward deeper comprehension.

Second, short and targeted micro-skill routines should be incorporated into every reading session. Teachers can adopt a "notice–name–apply" cycle to highlight conjunctions, transitions, punctuation and closing strategies in model texts before asking students to integrate these features into their own work. Vocabulary development should be supported through contextual clue identification, synonym–antonym associations and using new words in multiple contexts.

Third, structured and cognitively demanding prompts should be used to enhance inference making and thematic understanding. Guiding students to explore cause–effect relationships, character motivations and underlying messages encourages the development of higher-order reading skills that extend beyond literal recall.

Fourth, multimodal and scene-based activities can deepen students' detail awareness. Micro-drama, scene reconstruction, sequencing tasks and visual organisers such as storyboards or event maps enable learners to identify narrative structure more effectively and strengthen their ability to track key events.

Fifth, explicit teacher modelling is essential. Think-aloud demonstrations that show how to identify the main idea, extract themes, interpret character actions and follow the narrative structure provide students with a clear cognitive model they can emulate during independent reading.

Sixth, peer interaction should be strategically integrated to support motivation and comprehension. Collaborative retellings, peer discussions and evidence-based peer feedback routines can reinforce text understanding while building social engagement around reading tasks.

Finally, vocabulary development ought to be supported through repeated reading cycles and targeted instruction. AI-generated stories alone were not sufficient for fostering word consciousness; therefore, short repeated-reading routines and mini-glossary activities centred on important or unfamiliar words can help students internalise vocabulary more effectively. Encouraging students to integrate culturally meaningful elements into AI prompts can further strengthen comprehension by fostering personal and contextual relevance.

Conflicts of Interest: The authors declare that there is no conflict of interest.

Ethical Statement: Ethical approval for this research was granted by the Ethics Committee of the Graduate School of Çanakkale Onsekiz Mart University (decision date: 04/04/2024; approval number: 05/60; document reference: E-84026528-050.99-2400100807).

Credit Author Statement: The first author conducted data collection, analysis, reporting, and manuscript preparation. Academic supervision, editing and revisions were provided by the second author.

Generative AI Usage Statement: The authors declare that Generative AI tools were used to assist with refining language, generating alternative phrasing, ensuring grammatical accuracy, and improving overall readability. The authors have reviewed and edited the generated content to ensure its accuracy and alignment with their intended meaning. The authors take full responsibility for the content of the publication.

References

Ademola, O. E. (2024). Reading strategies in the AI age: Enhancing comprehension and engagement with advanced technologies. In *Proceedings of the 38th iSTEAMS Bespoke Conference* (pp. 105–124). Society for Multidisciplinary & Advanced Research Techniques (SMART). https://doi.org/10.22624/AIMS/ACCRABESPOKE2024P13

Akyol, H., & Çoban Sural, Ü. (2020). Developing Reading, Reading Comprehension and Motivation: An Action Research Study. *Education and Science*, *46*(205), 69-92. https://doi.org/10.15390/EB.2020.8977

Aşıkcan, M., & Saban, A. (2021). An action research on improving fluent reading skills of third-grade primary school students. *Education and Science*, *46*(205), 19–47. https://doi.org/10.15390/EB.2020.9015

Baker, L., & Wigfield, A. (1999). Dimensions of children's motivation for reading and their relations to reading activity and reading achievement. *Reading Research Quarterly, 34*(4), 452–477. https://doi.org/10.1598/RRQ.34.4.4

Borella, E., Carretti, B, & Pelegrina, S. (2010). The specific role of inhibition in reading comprehension in good and poor comprehenders. *Journal of Learning Disabilities*, *43*(6), 541–552.

Cain, K., & Oakhill, J. V. (1999). Inference making ability and its relation to comprehension failure in young children. *Reading and Writing*, *11*(5-6), 489–503. https://doi.org/10.1023/A:1008084120205

Chango García, A. P., & Lugmania Borja, K. D. (2025). *Didactic strategies based on ChatGPT to improve the reading skills of A2.2 level students* [Yayınlanmamış lisans araştırma raporu, Cotopaxi Teknik Üniversitesi, Ulusal ve Yabancı Diller Eğitimi Bölümü]. Pujilí.

Cruz Brotons, C. (2025). *Enhancing EFL literacy in upper primary education through a Generative Pre-Trained Transformer (GPT)* (Unpublished undergraduate thesis). University of Alicante, Faculty of Education, Spain.

Dessemontet, R. S., Morin, D., & De Chambrier, A.-F. (2024). Effects of repeated reading on fluency and comprehension: Evidence from elementary classrooms. *Journal of Educational Research*, *117*(3), 215–229.

Elmaadaway, M. A. N., El-Naggar, M. E., & Abouhashesh, M. R. I. (2025). Improving primary school students' oral reading fluency through voice chatbot-based AI. *Journal of Computer Assisted Learning*, *41*(3), e70019. https://doi.org/10.1111/jcal.70019

Esmer, B. (2019). *Sessiz ve sesli okuma akıcılığının, okuduğunu anlama, okuma benlik kavramı, okuma katılımı ve resimli öykü kitaplarına verilen tepkilerle doğrudan ve dolaylı ilişkileri* (Unpublished doctoral dissertation). Gazi University, Institute of Educational Sciences, Ankara, Turkey.

Fielding-Barnsley, R. (2000). Predicting future reading achievement from kindergarten measures: A follow-up study. *Australian Journal of Learning Disabilities*, *5*(2), 36–47.

Guthrie, J. T., & Wigfield, A. (2000). Engagement and motivation in reading. In M. L. Kamil, P. B. Mosenthal, P. D. Pearsonve R. Barr (Eds.), *Handbook of reading research* (Vol. 3, pp. 403–422). Lawrence Erlbaum Associates.

Henkel, O. (2025). Using large language models and speech-to-text models to facilitate the assessment of basic literacy in Ghana [Doctoral dissertation, University of Oxford].

Kintsch, W. (1998). *Comprehension: A Paradigm for Cognition*. Cambridge University Press.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Sage Publications.

Millî Eğitim Bakanlığı [MEB]. (2024). *Türkçe dersi öğretim programı (ilkokul ve ortaokul 1-8. sınıflar)*. Millî Eğitim Bakanlığı.

Morgan, P. L., & Fuchs, D. (2007). Is there a bidirectional relationship between children's reading skills and reading motivation? *Exceptional Children, 73*(2), 165–183.

Pereyra, A. E. (2025). Enhancing the reading comprehension skills of grade 4 pupils in selected schools in Catanauan I District through the use of pre-teaching vocabulary and repeated reading methods. *EPRA International Journal of Environmental Economics, Commerce and Educational Management*, *12*(7). https://doi.org/10.36713/epra0414

Rosenblatt, L. M. (1938). *Literature as exploration*. D. Appleton-Century.

Rosenblatt, L. M. (1978). *The reader, the text, the poem: The transactional theory of the literary work*. Carbondale, IL: Southern Illinois University Press.

Samuels, S. J. (1997). The method of repeated readings. *The Reading Teacher*, *50*(5), 376–381.

Schwanenflugel, P. J., Hamilton, A. M., Kuhn, M. R., Wisenbaker, J. M., & Stahl, S. A. (2004). Becoming a fluent reader: Reading skill and prosodic features in the oral reading of young readers. *Journal of Educational Psychology*, *96*(1), 119–129.

Whalley, K., & Hansen, J. (2006). The role of prosodic sensitivity in children's reading development. *Journal of Research in Reading*, *29*(3), 288–303.

Wigfield, A., & Guthrie, J. T. (1997). Motivation for reading: An overview. *Educational Psychologist, 32*(2), 57–58.

Yılmaz, M. (2009). Sesli okuma hatalarının tekrarlı okuma yöntemiyle düzeltilmesi. *Millî Eğitim Dergisi, 183*, 19–41.