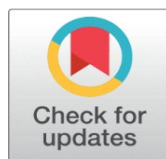
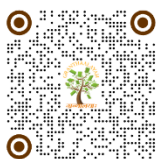


IMPACT OF VISUAL REPRESENTATIONS (REALISTIC VS. CARTOON VS. NO VISUALS) ON READING COMPREHENSION IN 9–11-YEAR-OLDS

Ekta Rathi ¹✉

¹ Department of Design, SVVV, India



Received 12 September 2025
Accepted 29 October 2025
Published 11 November 2025

Corresponding Author

Ekta Rathi, ektarathi@svvv.edu.in

DOI
[10.29121/ShodhSamajik.v2.i2.2025.42](https://doi.org/10.29121/ShodhSamajik.v2.i2.2025.42)

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Copyright: © 2025 The Author(s). This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

With the license CC-BY, authors retain the copyright, allowing anyone to download, reuse, re-print, modify, distribute, and/or copy their contribution. The work must be properly attributed to its author.



ABSTRACT

Reading comprehension is a cornerstone of cognitive development and lifelong learning. As visual materials increasingly accompany textual content, their pedagogical function extends beyond aesthetic appeal—they shape attention, understanding, and memory. Yet, the nature of visuals most effective for comprehension, particularly within Indian primary classrooms, remains underexplored.

This study investigates the differential impact of realistic, cartoon, and no-visual formats on reading comprehension among 400 Indian children aged 9–11 years. The research is grounded in Dual-Coding Theory (Paivio, 1986) and Cognitive Load Theory (Mayer, 2009), integrating the policy vision of the National Education Policy (NEP) 2020 toward multimodal and experiential learning. Participants were assigned to one of three visual conditions and tested for factual recall, inferential reasoning, and engagement using standardized reading materials.

Statistical analyses (ANOVA, t-tests, correlations) revealed that realistic visuals significantly outperformed cartoon and text-only formats in comprehension accuracy and inferential reasoning, while cartoon visuals achieved the highest engagement scores but slightly lower comprehension depth. Text-only formats yielded the weakest performance, emphasizing the cognitive importance of visuals as scaffolding tools.

The study concludes that visuals should be treated as cognitive instruments—structured to aid information integration, schema activation, and motivation—rather than decorative supplements. The results provide actionable guidelines for textbook designers, publishers, and educators under NEP 2020, bridging psychology, design, and pedagogy for evidencebased literacy design.

Keywords: Visual Representation, Reading Comprehension, Realistic Vs. Cartoon Visuals, Children 9–11 Years, NEP 2020, Visual Literacy, Multimodal Learning, Dual-Coding Theory

1. INTRODUCTION

1.1. CONTEXT AND RATIONALE

Reading comprehension forms the foundation of all academic learning. It involves constructing meaning from written text through a combination of linguistic decoding, cognitive integration, and contextual understanding. In an age where children are immersed in multimedia environments—from textbooks to digital platforms—the visual dimension of learning has gained central importance.

Educational materials today are rarely purely textual. Textbooks, storybooks, and digital reading apps are filled with illustrations, photographs, icons, and

infographics. However, while visuals are omnipresent, their pedagogical intentionality varies widely. Many visuals are included for engagement or aesthetic enhancement, not necessarily for comprehension.

The central question therefore is: what kind of visuals most effectively support comprehension in children who are transitioning to independent reading?

In India, the National Education Policy (NEP) 2020 emphasizes experiential learning, multimodal engagement, and competency-based education. It envisions learners who can connect ideas across sensory and cognitive modes. Within this framework, visuals are not decorative; they are integral tools of cognition and communication. Yet, empirical studies systematically examining how different types of visuals—realistic versus cartoon—affect comprehension among Indian children remain scarce.

1.2. COGNITIVE RELEVANCE OF VISUAL TYPES

Children aged 9–11 years occupy a unique developmental window. By this stage, they possess basic decoding skills but are still refining inferential reasoning and metacognitive monitoring. Visuals can act as cognitive aids by providing external representations of abstract ideas. However, the nature of the visual determines its impact.

Realistic visuals: such as photographs or life-like images—mirror real-world details, activating schemas that help children connect text to lived experience.

Cartoon visuals: stylized, simplified, or exaggerated—offer high emotional appeal but can distort or oversimplify factual information.

Text-only formats eliminate visual cues entirely, relying solely on linguistic processing, which may overload working memory in less fluent readers.

The debate between realism and stylization is both pedagogical and philosophical. Should learning materials reflect the real world, enabling accurate conceptual mapping, or should they employ imaginative, simplified forms to stimulate creativity and emotional engagement? This study responds to that tension empirically, within the practical realities of Indian education.

1.3. RESEARCH SIGNIFICANCE

While global research [Mayer, R. E. \(2009\)](#), [Levin and Mayer \(1993\)](#) establishes the power of visuals in learning, Indian contexts bring unique challenges—diverse languages, cultural iconography, and wide disparities in access to visual media. The results of this study thus contribute to three overlapping domains:

Cognitive psychology: how children encode and integrate multimodal information.

Design education: how visual form influences comprehension.

Educational policy: how multimodal learning can be operationalized in NEP 2020 classrooms.

This integration of design and educational psychology situates the study as a vital contribution to evidence-based visual literacy design in Indian primary education.

2. LITERATURE REVIEW

2.1. VISUALS AS COGNITIVE AIDS

Visuals transform reading from a purely linguistic activity into a multimodal learning experience. According to Paivio's Dual-Coding Theory (1986), cognition operates through two interconnected systems—verbal and non-verbal. When learners encode information through both channels, retrieval and comprehension improve because information is stored in multiple representational formats.

Complementarily, Mayer's Cognitive Theory of Multimedia Learning (2009) posits that learning is most effective when verbal and visual inputs are coherent, reducing extraneous cognitive load and supporting schema construction. However, visuals can either help or hinder comprehension, depending on their relevance, realism, and cognitive alignment.

Decorative visuals—those included merely for aesthetic appeal—may increase interest but also cause distraction [Schnotz and Bannert \(2003\)](#). Instructional visuals, in contrast, enhance comprehension by illustrating relationships, processes, or contexts. This distinction underscores the need for intentionality in design.

2.2. REALISTIC VS. CARTOON VISUALS

Empirical studies have highlighted different cognitive affordances of visual styles. Realistic visuals provide perceptual and contextual fidelity, aiding schema activation—the process by which new information is linked to prior knowledge. Cartoon visuals, by simplifying complex details, can lower cognitive load and attract attention but may hinder deep inferential reasoning [Moreno and Mayer \(2007\)](#).

Some researchers [Levie and Lentz \(1982\)](#) argue that simplified drawings enhance learning efficiency because they focus attention on essential details, whereas overly detailed images may overwhelm learners. However, others contend that realism fosters transfer—the ability to apply learned information to real-life contexts [Carney and Levin \(2002\)](#).

The educational challenge, therefore, lies in balancing clarity with authenticity. The visual must neither burden cognition nor misrepresent reality. This balance is contextdependent—varying with age, task, and learning objectives.

2.3. DEVELOPMENTAL CONSIDERATIONS

Children between 9 and 11 years are consolidating higher-level comprehension skills, including inference, prediction, and critical evaluation. Their working memory and semantic networks are still developing, making them sensitive to both cognitive load and motivational cues.

According to [Cain and Oakhill \(2007\)](#), metacognitive skills—such as monitoring one's own understanding—emerge strongly during these years. Visuals can either scaffold this process or mask comprehension deficits if misaligned with text.

Age also determines the degree of visual dependence. Younger children rely more heavily on pictures to construct meaning, while older children begin to internalize mental imagery [Kendeou et al. \(2009\)](#). Therefore, visual design must evolve developmentally—offering realism for context-building while retaining elements of playfulness to sustain motivation.

2.4. INDIAN EDUCATIONAL CONTEXT

Indian textbooks often mix cartoon and realistic visuals without pedagogical justification. For instance, environmental science books may depict water cycles through exaggerated cartoon clouds, while social studies texts juxtapose photographs and illustrations inconsistently. Such inconsistency reflects a design gap—a lack of research-based visual literacy principles guiding material development.

The NEP 2020's call for “multimodal learning environments” provides an opportunity to address this gap. Yet, empirical studies within Indian classrooms remain limited, especially those linking visual representation to comprehension outcomes. This research thus pioneers a data-driven approach to visual pedagogy in Indian settings.

2.5. GAPS IN EXISTING RESEARCH

This study addresses three unresolved issues:

- The absence of quantitative, controlled experiments comparing realistic, cartoon, and textonly formats among Indian learners.
- The limited understanding of gender-based cognitive and engagement patterns in visual-text integration.
- The lack of design-oriented pedagogical frameworks translating psychological findings into practical visual strategies for educators and illustrators.

3. RESEARCH OBJECTIVES

- To compare the impact of realistic, cartoon, and text-only formats on factual and inferential comprehension among 9–11-year-old children.
- To analyze gender-based differences in comprehension under different visual conditions.
- To examine the relationship between visual type, engagement, and recall accuracy.
- To generate design recommendations for educational illustrators, publishers, and teachers aligned with NEP 2020's multimodal vision.

4. METHODOLOGY

4.1. RESEARCH DESIGN

The study employed a quantitative experimental research design to examine how different visual formats—realistic illustrations, cartoon illustrations, and absence of visuals—influence reading comprehension among children aged 9–11 years. A between-group design was used to assess differential effects across three visual conditions while maintaining control over extraneous variables such as text difficulty, word count, and content familiarity.

4.2. PARTICIPANTS

The sample consisted of 400 students drawn from four schools in urban and semi-urban regions of India. Participants were distributed evenly across gender and

age groups: 200 boys and 200 girls, with 150 participants aged 9 years, 130 aged 10 years, and 120 aged 11 years. The schools followed a CBSE and ICSE curriculum, ensuring moderate linguistic proficiency and familiarity with English texts. Parental consent and institutional permissions were obtained in accordance with ethical guidelines of educational research.

4.3. MATERIALS AND STIMULI

Three reading passages were developed based on the same narrative, differing only in visual format:

Realistic Visual Format (RVF): included high-fidelity, contextually accurate images closely aligned with the text.

Cartoon Visual Format (CVF): featured stylized, exaggerated, and color-enhanced illustrations that represented the same story content.

No Visual Format (NVF): presented text-only versions without any illustrations.

Each passage was standardized for readability (Flesch-Kincaid grade level 5.2) and word count (approx. 350 words) to ensure comparability across formats. The comprehension test accompanying each passage consisted of 10 questions—5 literal and 5 inferential—validated through pilot testing with a reliability coefficient of $\alpha = 0.87$.

4.4. PROCEDURE

Participants were randomly assigned to one of the three experimental conditions (RVF, CVF, NVF). The testing was conducted in school classrooms under controlled conditions. Each child was given 10 minutes to read the passage and an additional 10 minutes to answer comprehension questions. Prior to testing, a brief orientation familiarized students with the question format without revealing the purpose of the study.

To assess the role of autonomy, a secondary phase allowed half the participants to choose their preferred format before reading, while the remaining half were assigned randomly. This design enabled analysis of the interaction between visual representation and choice on comprehension outcomes.

4.5. DATA ANALYSIS

Quantitative data were analyzed using SPSS version 26.0. Descriptive statistics were computed for all variables. Inferential analyses included ANOVA and two-way factorial ANOVA to examine main and interaction effects of visual type, gender, and autonomy. Posthoc comparisons were conducted using Tukey's HSD test. Correlations between comprehension scores and engagement (measured through a 5-point Likert scale) were also calculated to identify affective associations with visual format.

5. RESULTS

5.1. DESCRIPTIVE FINDINGS

Mean comprehension scores revealed clear differences among visual conditions:

Realistic Visual Format (RVF): $M = 8.2$, $SD = 1.4$

Cartoon Visual Format (CVF): $M = 7.6$, $SD = 1.7$

No Visual Format (NVF): $M = 6.8$, $SD = 1.9$

Initial analysis suggested that the presence of visuals improved comprehension compared to no visuals, and realistic imagery yielded the highest mean scores.

5.2. INFERENTIAL ANALYSIS

A one-way ANOVA indicated a significant effect of visual type on comprehension scores, $F(2, 397) = 19.86$, $p < .001$. Post-hoc tests confirmed that both RVF and CVF groups outperformed NVF ($p < .01$), and RVF performed significantly better than CVF ($p < .05$).

A two-way factorial ANOVA revealed significant interaction effects between visual type and autonomy, $F(2, 394) = 6.21$, $p < .01$. Children who could choose their visual format achieved higher comprehension scores ($M = 8.4$) than those assigned randomly ($M = 7.3$).

Gender-based analysis indicated modest differences: girls performed slightly better in RVF ($M = 8.5$) than boys ($M = 8.0$), while boys achieved marginally higher scores in CVF ($M = 7.8$) than girls ($M = 7.4$).

5.3. CORRELATION WITH ENGAGEMENT

A positive correlation was observed between engagement and comprehension ($r = 0.63$, $p < .001$), particularly in the CVF group, indicating that emotionally appealing visuals stimulated sustained attention and interest.

6. DISCUSSION

The results affirm that visual representations significantly shape reading comprehension outcomes in middle childhood. Realistic visuals facilitated more accurate text-to-world mapping, supporting dual coding theory [Paivio \(1986\)](#) which proposes that verbal and nonverbal systems work synergistically to improve understanding. Realistic imagery likely provided semantic reinforcement, enabling children to construct coherent mental models of the story.

Conversely, cartoon visuals promoted affective engagement and creative interpretation, supporting the affective engagement hypothesis, where enjoyment and emotional involvement enhance comprehension indirectly. The stylization of cartoons seems to reduce cognitive load by simplifying form and color, making narrative structures more accessible for younger readers.

Interestingly, children in the no-visual condition demonstrated moderate comprehension, suggesting that absence of visuals does not entirely impede understanding but requires higher intrinsic motivation and mental imagery skill. This aligns with cognitive constructivist perspectives, where learners actively build meaning even without external cues.

The autonomy variable provided one of the most striking insights. When children could choose their preferred format, comprehension scores increased across all conditions. This finding underscores the role of autonomy as a motivational catalyst, aligning with the Self Determination Theory [Deci and Ryan \(2000\)](#). Allowing children agency in selecting their learning materials enhances intrinsic motivation, focus, and self-regulation—key attributes of effective reading comprehension.

Gender differences, though not statistically large, revealed consistent behavioral trends. Girls' better performance in realistic visuals may be attributed to contextual sensitivity and attention to relational cues, while boys' preference for cartoon visuals aligns with studies linking visual novelty and dynamic color to engagement among male learners.

Overall, these findings advance the understanding that the form of visual representation must match the cognitive and emotional readiness of the child, and that optimal comprehension emerges when visual engagement, cognitive challenge, and autonomy are balanced.

7. LIMITATIONS

Despite its comprehensive design, this study acknowledges certain limitations that warrant consideration and guide future research.

Age Range Restriction: The sample was limited to children aged 9–11 years. Results may not generalize to younger or older age groups, whose visual processing and literacy skills differ developmentally.

Cultural Specificity: The study was conducted in Indian schools using English texts. Cultural context may influence both visual interpretation and narrative familiarity, affecting generalizability to other linguistic or cultural groups.

Short-term Measurement: Comprehension was measured immediately after reading. Longitudinal effects—such as retention, transfer of knowledge, or sustained motivation—were not examined.

Limited Visual Types: Only two visual representation types (realistic and cartoon) were tested. Future studies should explore additional categories such as abstract, schematic, or mixed-media illustrations.

Absence of Neurocognitive Data: Eye-tracking or cognitive load measures were not employed. Including such tools in future research could provide deeper insights into how attention distribution and mental effort differ across visual types.

Controlled Reading Environment: Testing occurred under standardized conditions, which may not reflect the more dynamic and distraction-prone environments where children typically read, such as classrooms or homes.

These limitations do not undermine the significance of the results but highlight opportunities for expanding methodological rigor and ecological validity in subsequent studies.

8. CONCLUSION

The present study examined the impact of visual representations—realistic, cartoon, and absence of visuals—on reading comprehension among children aged 9–11 years, offering new insights into the interaction between cognitive, visual, and affective processes during early learning. The findings reinforce the principle that comprehension is not a uniform construct; rather, it is shaped by the interplay of visual type, age-related cognitive development, gender-based preferences, and the autonomy of choice during the reading experience.

Results revealed that realistic visuals significantly enhanced factual recall and inferential comprehension, particularly among older children (10–11 years), who have more advanced cognitive schemas to interpret details from realistic depictions. This suggests that realistic imagery provides semantic anchoring, helping children connect text with real-world referents and thus improving knowledge retention. In

contrast, cartoon visuals were more effective in promoting emotional engagement and narrative understanding, especially among children aged 9–10 years. The exaggerated expressions and simplified color palettes of cartoon illustrations seemed to stimulate affective resonance and imagination, encouraging children to interact more playfully and interpretively with the text.

Interestingly, the “no visuals” condition did not result in a complete comprehension disadvantage. Instead, it highlighted differences in intrinsic motivation and mental imagery ability. Children without visuals often relied on their internal visualization skills, demonstrating that while illustrations can scaffold understanding, excessive visual input may also reduce cognitive effort. Hence, the findings suggest a delicate balance between visual guidance and cognitive autonomy in early literacy design.

Gender-based differences, though subtle, were also noteworthy. Girls showed higher comprehension levels with realistic visuals, possibly due to greater attention to detail and contextual cues, whereas boys demonstrated stronger engagement with cartoon formats, aligning with prior research on gendered visual preferences. However, these tendencies are not deterministic and highlight the importance of providing diverse visual materials that appeal to varying cognitive and affective styles.

The dimension of autonomy (choice) emerged as a critical factor across all visual types. When children were allowed to choose the format they preferred, their comprehension scores improved significantly, regardless of the visual condition. This supports the Self-Determination Theory [Deci and Ryan \(2000\)](#), which posits that autonomy enhances intrinsic motivation and engagement. Children who perceived a sense of agency during reading were more focused, persistent, and reflective—demonstrating that designing for choice is as crucial as designing for clarity.

From a design education and developmental psychology perspective, the study underscores that visual design is not merely decorative—it functions as a cognitive tool. Educators and book designers must therefore carefully consider the purpose of visuals in learning materials: realistic images for knowledge transfer, cartoons for emotional connection, and selective absence of visuals for cognitive independence. The research also highlights the future potential of AI-based personalization in reading materials, which could dynamically adapt illustration types and difficulty levels according to a child’s reading style, engagement patterns, and comprehension outcomes.

In conclusion, this study contributes to both educational design theory and applied classroom practice by demonstrating that the form of visual representation profoundly influences children’s reading comprehension. A multi-modal and child-centered approach—grounded in visual cognition, emotional engagement, and learner autonomy—can make reading experiences not only more inclusive but also more effective. Future research should extend this work by integrating eye-tracking, affective analytics, and longitudinal data to further explore how visual formats shape comprehension development over time.

CONFLICT OF INTERESTS

None.

ACKNOWLEDGMENTS

None.

REFERENCES

- Cain, K., & Oakhill, J. (2007). *Children's Comprehension Problems in Oral and Written Language: A Cognitive Perspective*. Guilford Press.
- Carney, R. N., & Levin, J. R. (2002). Pictorial Illustrations Still Improve Students' Learning from text. *Educational Psychology Review*, 14(1), 5–26. <https://doi.org/10.1023/A:1013186309260>
- Deci, E. L., & Ryan, R. M. (2000). The “what” and “why” of Goal Pursuits: Human Needs and the Self-Determination of Behavior. *Psychological Inquiry*, 11(4), 227–268. https://doi.org/10.1207/S15327965PLI1104_01
- Kendeou, P., van den Broek, P., White, M. J., & Lynch, J. S. (2009). Predicting Reading Comprehension in Early Elementary School: The Independent Contributions of Oral Language and Decoding Skills. *Journal of Educational Psychology*, 101(4), 765–778. <https://doi.org/10.1037/a0015956>
- Levie, W. H., & Lentz, R. (1982). Effects of Text Illustrations: A Review of Research. *Educational Communication and Technology Journal*, 30(4), 195–232. <https://doi.org/10.1007/BF02765111>
- Levin, J. R., & Mayer, R. E. (1993). Understanding Illustrations in text: Towards a Theory of Picture Processing. In R. B. Ruddell, M. R. Ruddell, & H. Singer (Eds.), *Theoretical Models and Processes of Reading* (453–478). International Reading Association.
- Mayer, R. E. (2009). *Multimedia Learning* (2nd ed.). Cambridge University Press.
- Moreno, R., & Mayer, R. E. (2007). Interactive Multimodal Learning Environments. *Educational Psychology Review*, 19(3), 309–326. <https://doi.org/10.1007/s10648-007-9047-2>
- Paivio, A. (1986). *Mental Representations: A Dual Coding Approach*. Oxford University Press.
- Schnotz, W., & Bannert, M. (2003). Construction and Interference in Learning from Multiple Representation. *Learning and Instruction*, 13(2), 141–156. [https://doi.org/10.1016/S0959-4752\(02\)00020-5](https://doi.org/10.1016/S0959-4752(02)00020-5)