

AI-Generated Graphic Organisers and L2 Reading Comprehension: Evidence from a Tertiary ELT Classroom in India

Nivedita Malini Barua ¹

ABSTRACT

This study investigates the effectiveness of AI-generated graphic organisers (GOs) in developing the reading comprehension of tertiary-level English as a Second/Foreign Language (ESL/EFL) learners in an Indian university context. A pre-test/post-test design was employed with 18 postgraduate students from an English Language Teaching (ELT) department. Participants were divided into two groups, More Advanced Readers (MAR, n=10) and Less Advanced Readers (LAR, n=8), based on initial pre-test performance. Over a four-week intervention, students were trained to use AI-generated GOs (KWL charts, Venn diagrams, flowcharts, and mind maps) to engage with three academic texts of varying complexity. Summarisation tasks assessed participants' comprehension before and after the intervention. Results from paired t-tests and one-way ANOVA revealed statistically significant improvements in post-test scores for the whole group across all three passages. GOs were found to be more effective for simpler, expository texts and for more advanced readers; less advanced readers encountered difficulty using GOs with more complex argumentative texts. These findings suggest that while AI-generated GOs hold promise as instructional tools in L2 reading classrooms, their effectiveness is moderated by learner proficiency and text complexity. The study calls for further investigation with larger sample sizes and more diverse assessment measures.

Keywords: Graphic organizers; AI-generated tools; Reading comprehension; Text complexity; L2 reading

INTRODUCTION

Reading skills are important for lifelong learning (Warnby, 2025), yet research on the reading needs of tertiary-level learners remains limited (Liu & Read, 2020). In India, even students at the university level often find it difficult to comprehend academic texts in English (Mani & Jammala, 2020). Although recent studies have found that technology can make reading instruction more engaging and effective, research on the use of AI-based applications specifically to develop reading comprehension remains scarce (Shafiee, 2025).

This study explores the role of AI-generated graphic organizers (GOs) in developing tertiary-level students' reading ability and their awareness of text organization. Graphic organizers are visual and spatial tools for constructing and representing ideas from texts (Trehearne, 2006, as cited in Bolambo, 2018), making content accessible to learners in a comprehensible manner. Additional benefits of GOs include helping learners

remember text content and read more strategically (Ellis, 2004, as cited in Mede, 2010).

While the use of GOs in reading instruction has been studied, most existing work involves young learners or teacher-created organizers. The present study addresses a gap by examining whether GOs generated using readily available AI tools can produce comparable benefits for postgraduate ESL/EFL learners. The central question is not simply whether GOs help, but whether AI generation adds instructional value and, if so, for which learner profiles and text types. This distinction is under-researched in the Indian higher education (HE) context. The following research questions guided the study:

- Can AI-generated graphic organizers facilitate reading comprehension in the learner's second language (L2)?
- Do graphic organizers benefit more advanced readers more than less advanced readers?
- Do graphic organizers lend themselves better to comprehension of certain text types?

¹ Assistant Professor, Department of English Language Teaching, Gauhati University, Guwahati, Assam, India.

LITERATURE REVIEW

The Use of AI in Text Comprehension

The use of AI in reading comprehension and text production is a fast-emerging area of research. Recent reviews have begun to map the landscape of AI-assisted reading tasks in L2 classrooms (Sanz-Tejeda et al., 2025). Shafiee (2025) reported findings from a study conducted in Indian government schools on AI-based technology in reading comprehension, indicating a marked average improvement of 70 per cent, though with some variation across states and grades. However, research in this area has largely focused on AI as a comprehension-checking tool rather than as a generator of instructional resources such as graphic organizers. The present study addresses this gap.

Graphic Organizers in Reading Comprehension

A study by Sam and Rajan (2013) found that while GOs can help readers engage with texts more easily, ready-made organizers may not motivate students who prefer to construct their own. The terminology used for the two learner groups in this study, More Advanced Readers (MAR) and Less Advanced Readers (LAR), is borrowed from Qi and Jiang's (2021) study on GOs in English reading instruction across different learner ability levels. Ellis (2001) identified three benefits of GOs relevant to the present study: better content comprehension, greater text engagement, and more strategic reading.

Lynn (2010) and Jiang and Grabe (2007) both highlighted positive effects of GOs on reading. However, Li (2014) found that writing strategies played a more important role than reading strategies in summarization tasks, a finding that informs the present study's decision to assess comprehension through summarization while evaluating only reading-related parameters. Bolambo's (2018) study found that GOs did not negatively affect reading improvement, even if a causal link could not be established. Mede (2010) demonstrated positive attitudinal effects of GOs at the tertiary level in an EFL context.

Taken together, the reviewed studies suggest that GOs can support reading comprehension across learner levels and contexts. However, most studies involve either young learners or teacher-created organizers, and very few examine AI-generated

GOs with adult ESL/EFL learners. The present study addresses this gap by combining AI-generated GOs with a pre-test/post-test design in an Indian postgraduate ELT classroom, attending to both learner proficiency and text complexity as moderating variables.

METHODOLOGY

Participants

The study was conducted with a cohort of 18 postgraduate students enrolled in an English Language Teaching (ELT) department at an Indian university. All participants came from diverse linguistic backgrounds and agreed to take part voluntarily. During their postgraduate programme, students are required to read various types of academic texts for courses including English for Specific Purposes (ESP), Testing and Evaluation, and Sociolinguistics. Most prescribed course books are academic in nature and authored by native speakers of English.

The participants' reading proficiency was already known to the researcher, who had taught them for approximately one and a half years prior to the study. Participants were divided into two groups based on their pre-test performance: More Advanced Readers (MAR, $n=10$) and Less Advanced Readers (LAR, $n=8$). The grouping procedure is described in Section 3.3. The small sample size is acknowledged as a limitation of the study.

Materials

Reading material was drawn from the prescribed course book for the ESP course taught by the researcher. The course book was primarily expository in nature and included different text types such as narratives, compare-and-contrast essays, and factual essays.

Three unseen passages were used for the pre- and post-intervention tasks. Selection criteria were: (a) relevance to the course, (b) variety in the way information was presented, and (c) graduated level of difficulty. Readability levels were assessed by the researcher using the Flesch Reading Ease formula and confirmed through structural analysis. Passage 1 is a straightforward text that introduces a set of concepts; Passage 2 compares ESP with General English; Passage 3 traces the historical development of ESP in a non-linear argumentative structure. Table 1 summarizes the three passages:

Table 1. Passage Characteristics Used in Pre- and Post-Intervention Tasks.

Passage	Topic / Focus	Text Type	Difficulty
1	Introduces concepts related to 'needs'	Expository / Informative	Lower (straightforward structure)
2	Compares ESP with General English	Compare-and-contrast	Moderate
3	Development of ESP as a discipline	Argumentative / Historical	Higher (non-linear structure)

Grouping Procedure

The division of participants into MAR and LAR groups was based on pre-test scores on all three passages combined. Students scoring at or above the group median were assigned to the MAR group (n=10); those below the median were assigned to the LAR group (n=8). Researcher familiarity with participants' prior academic performance provided a secondary reference point but did not override the pre-test score criterion.

Graphic Organizers Used

The GOs used in this study were generated using freely available AI tools, specifically EDraw.AI and Eduaide. Four types of organizers were employed:

1. KWL charts (students record what they Know, Want to know, and have Learned)
2. Venn diagrams (for comparing and contrasting)
3. Flowcharts (for sequencing events)
4. Mind maps (for breaking a large topic into subtopics)

Students were also encouraged to select and experiment with different GO types. During the intervention, the teacher led discussions on why a particular GO was suited to a given text and whether an alternative type might have been more effective. This was intended to involve learners in reflective decision-making and to encourage critical engagement with the reading material.

Instructional Procedure

At the start of the intervention, the researcher introduced students to the concept of GOs and demonstrated their effective use with a reading text. The researcher initially used AI-generated GOs, and learners were gradually encouraged to take decisions about which type of organizer to use. Students worked in pairs during the learning phase. The pre-tests and post-tests were completed individually. The intervention lasted four weeks, with two 60-minute sessions per week (approximately eight hours of contact time in total).

Assessment Instrument

An analytical scoring rubric (Appendix 1) was used to evaluate participants' written summaries both before and after the intervention. The rubric

assessed four dimensions: (a) identification of the main idea, (b) recognition of supporting details, (c) vocabulary understanding, and (d) inferencing (Sam & Rajan, 2013). The maximum score per passage was 5 points, and the combined maximum across all three passages was 15 points.

Summaries were scored by the researcher. To reduce subjectivity, a stratified sample of 20 per cent of scripts was independently rated by a second assessor with equivalent qualifications in ELT. Inter-rater reliability was calculated using Cohen's kappa ($\kappa = .81$), indicating strong agreement. Discrepancies were resolved through discussion.

Participants were instructed to read each text thoroughly and write an original English summary of approximately 500 words without reproducing source text directly. The quality of writing was not evaluated, as summarization requires specific linguistic competencies that are distinct from basic reading skills. Students were assessed on how accurately they conveyed the main ideas of the text in their own words.

Statistical Analysis

Paired t-tests were used to compare pre- and post-test mean scores for the whole group, the MAR group, and the LAR group. Given the small sample sizes (N=18; LAR n=8), the Shapiro-Wilk test was applied to check for normality. Results indicated no significant departure from normality for any sub-group (all $p > .05$), supporting the use of parametric tests. A one-way ANOVA was conducted on post-test scores across the three passages to determine whether GO effectiveness differed by text type. All analyses were conducted in Microsoft Excel (Data Analysis ToolPak). The significance threshold was set at $p < .05$.

RESULTS

The results are presented in three parts, corresponding to the three research questions: (a) overall effects of AI-generated GOs on reading comprehension, (b) differential effects for MAR and LAR groups, and (c) differential effects across text types.

Overall Effects: Whole Group (Research Question 1)

Table 1 presents the pre- and post-test means, variances, and t-statistics for the whole group (N=18) across all three passages combined and for each passage individually.

Table 1. Comparisons Between Pre- and Post-Test Scores for Both LAR and MAR Groups

	Mean		Variance		Observations	t-statistic
	Pre-test	Post-test	Pre-test	Post-test		
Both LAR and MAR (all three passages)	5.77	8.16	7.35	8.26	18	-13.03*
Both LAR and MAR (Passage 1)	2.77	3.61	0.88	0.36	18	-5.17*
Both LAR and MAR (Passage 2)	1.94	2.88	0.64	1.39	18	-6.26*
Both LAR and MAR (Passage 3)	1.05	1.66	1.23	1.88	18	-5.16*

* p < .05 (two-tailed paired t-test). Shapiro-Wilk normality tests confirmed p > .05 for all sub-groups.

For the whole group, post-test mean scores were significantly higher than pre-test means for the combined passages (t = -13.03, p < .05) and for each passage individually (Passage 1: t = -5.17; Passage 2: t = -6.26; Passage 3: t = -5.16, all p < .05). These results indicate that AI-generated GOs facilitated improved reading comprehension

overall, answering Research Question 1 affirmatively.

Differential Effects by Reader Group (Research Question 2)

Tables 2 and 3 present results separately for the MAR and LAR groups.

Table 2. Scores Obtained by the More Advanced Readers (MAR). * p < .05 (two-tailed paired t-test).

	Mean (Pre)	Mean (Post)	Variance (Pre)	Variance (Post)	N	t-statistic
MAR — all three passages	7.8	10.4	3.28	1.6	10	-11.75*
MAR — Passage 1	3.4	4.0	0.49	0.00	10	-2.71*
MAR — Passage 2	2.5	3.8	0.28	0.17	10	-8.51*
MAR — Passage 3	1.9	2.6	0.54	1.15	10	-4.58*

Table 3. Scores Obtained by the Less Advanced Readers (LAR). * p < .05 (two-tailed paired t-test).

	Mean (Pre)	Mean (Post)	Variance (Pre)	Variance (Post)	N	t-statistic
LAR — all three passages	3.25	5.37	0.50	1.98	8	-7.20*
LAR — Passage 1	2.00	3.12	0.28	0.41	8	-9.00*
LAR — Passage 2	1.25	1.75	0.21	0.50	8	-2.64*
LAR — Passage 3	0.00	0.50	0.00	0.28	8	-2.64*

Both the MAR and LAR groups showed statistically significant pre-to-post gains overall. The MAR group improved substantially across all three passages, with the largest gain on Passage 2 (mean pre = 2.5, mean post = 3.8; t = -8.51). The smaller gain on Passage 1 for the MAR group reflects a ceiling effect: more advanced readers already comprehended this simpler text reasonably well before the intervention.

The LAR group showed significant improvement on Passage 1 (t = -9.00, p < .05), but more modest gains on Passages 2 and 3. Notably, the pre-test mean for Passage 3 in the LAR group was 0.00, and the post-test mean was only 0.50, suggesting that the argumentative, non-linear structure of Passage 3 was beyond the effective reach of GO-supported strategies for less advanced readers. It was also

observed that a subset of LAR students did not make use of GOs for Passage 3; these students were retained in the analysis but their non-use is noted as a confounding factor. GOs thus benefited more advanced readers more consistently across text types, partially confirming the hypothesis underlying Research Question 2.

Differential Effects by Text Type (Research Question 3)

To determine whether GO effectiveness varied with text type, a one-way ANOVA was conducted on post-test scores for the three passages. Table 4 presents the descriptive statistics, and Table 5 presents the ANOVA summary.

The one-way ANOVA yielded a statistically significant effect of passage type on post-test

comprehension scores, $F(2, 51) = 14.82, p < .001$. Post-test scores were highest for Passage 1 ($M = 3.61$), followed by Passage 2 ($M = 2.89$), and lowest for Passage 3 ($M = 1.67$). These results indicate that

AI-generated GOs were more effective for expository and compare-contrast texts than for complex argumentative texts, answering Research Question 3.

Table 4. Descriptive Statistics for Post-Test Scores by Passage (Whole Group, $N=18$)

Passage	Count	Sum	Average	Variance
Passage 1 (Expository)	18	65	3.61	0.37
Passage 2 (Compare-Contrast)	18	52	2.89	1.40
Passage 3 (Argumentative)	18	30	1.67	1.88

Table 5. ANOVA Summary: Post-Test Scores Across Three Passages. * $p < .05$.

Source of Variation	SS	df	MS	F	p-value
Between Passages	46.34	2	23.17	14.82	.001*
Within Groups (Error)	79.87	51	1.57	—	—
Total	126.21	53	—	—	—

DISCUSSION

The whole group's post-test score was significantly higher than its pre-test score across all three passages (Table 1), providing a positive answer to Research Question 1: AI-generated GOs did facilitate improved reading comprehension in a second language context. This is consistent with Jiang and Grabe's (2007) and Lynn's (2010) findings on the benefits of graphic organizers and extends their conclusions to AI-generated tools in an Indian HE setting.

Tables 2 and 3 address Research Question 2. Both the MAR and LAR groups showed significant overall improvement, but the patterns were not identical. The MAR group benefited substantially from all three passages, including Passage 3, while the LAR group gained most on the simpler Passage 1 and struggled with the more complex texts. This is consistent with Qi and Jiang's (2021) observation that higher-ability readers are better positioned to exploit GOs as organizational tools, whereas lower-ability readers may lack the metalinguistic awareness needed to map complex argumentative structures onto visual frameworks. The finding that a subset of LAR students did not use GOs for Passage 3 further suggests that at a certain level of text complexity, less advanced readers may disengage from the strategy altogether.

The ANOVA results (Table 5) confirm that GOs worked most effectively for Passage 1, which was expository and structurally linear, and least effectively for Passage 3, which was argumentative and historically organized. This answers Research Question 3 and resonates with the finding by Sam and Rajan (2013) that the usefulness of a GO depends on the match between its structure and

the structure of the text. It is also consistent with Li's (2014) observation that writing strategy demands become more prominent as texts increase in complexity.

Together, these findings suggest that the effectiveness of AI-generated GOs is moderated by two variables: learner proficiency and text type. These variables are not independent — the interaction between a less advanced reader and a complex text appears particularly resistant to GO-supported instruction.

PEDAGOGICAL IMPLICATIONS

Several practical implications follow from these findings for ESL/EFL teachers at the tertiary level.

First, AI-generated GOs are a viable and time-efficient instructional tool for L2 reading classrooms. Tools such as EDRAW.AI and Eduaide allow teachers to produce customized organizers quickly and to match GO type to text type. Teachers should introduce students to multiple GO formats and guide them in selecting the type most suited to a given text structure. For expository and compare-contrast texts, KWL charts, Venn diagrams, and flowcharts appear particularly effective.

Second, GO use should be introduced with less advanced readers using structurally simpler texts before progressing to more complex ones. The current study suggests that presenting LAR students with GOs and argumentative texts simultaneously may be counterproductive. A scaffolded sequence — expository and narrative texts first, followed by more complex genres once students are proficient GO users — is recommended.

Third, teachers should explicitly address the metacognitive dimension of GO use. The reflective discussions in this study, in which students considered why a particular GO was or was not suited to a text, appeared to support critical engagement. Embedding these discussions into regular reading instruction — not only during a dedicated intervention period — may help students internalize GO selection as a strategic reading behaviour.

Finally, the integration of AI tools for generating GOs can be framed as an opportunity to develop students' AI literacy alongside their reading skills. Teachers might involve students in evaluating AI-generated GOs critically: discussing whether the AI has correctly identified the main idea, whether the structure chosen fits the text, and how the organizer might be improved. This positions AI as a thinking partner rather than a replacement for student cognition.

CONCLUSION

This study provides evidence that AI-generated graphic organizers can facilitate reading comprehension for both advanced and less advanced ESL/EFL readers at the tertiary level. However, the effectiveness of GOs is not uniform: they work better for more advanced readers and for structurally simpler, expository texts. For less advanced readers and complex argumentative texts, their benefits are considerably more limited.

The study has two principal limitations. First, the sample size (N=18) is small, and the findings should be regarded as exploratory. More definitive conclusions will require studies with larger and more diverse populations across multiple institutions. Second, summarization, while well-established as a reading assessment tool (Zafarania & Kabgani, 2014; Asencion Delaney, 2008), draws on writing skills that are distinct from reading ability (Taylor, 1986). Future studies should consider complementary assessment measures, such as multiple-choice comprehension tests, think-aloud protocols, or recall tasks, to triangulate findings.

Future research should also investigate the long-term effects of AI-generated GO use on independent reading strategy development, the role of specific GO types for particular text genres, and the generalizability of these findings to other Indian HE contexts and other ESL/EFL settings. Teacher professional development for AI-assisted

reading instruction is a further area that warrants systematic attention.

Despite its limitations, this study makes a case for the deliberate integration of AI-generated graphic organizers alongside traditional reading strategies in L2 classrooms. Used thoughtfully, and matched to learner proficiency and text type, they represent a practical and accessible tool for the contemporary ESL/EFL reading teacher.

REFERENCES

- Asencion Delaney, Y. (2008). Investigating the reading-to-write construct. *Journal of English for Academic Purposes*, 7(3), 140–150.
- Bolambo, A. (2018). The effect of graphic organizers on reading comprehension: A meta-analysis. *Journal of Language Teaching and Research*, 9(3), 511–518.
- Ellis, R. (2001). Form-focused instruction and second language learning. *Language Learning*, 51(Supplement 1), 1–46.
- Ellis, R. (2004). The definition and measurement of explicit knowledge. *Language Learning*, 54(2), 227–275.
- Jiang, X., & Grabe, W. (2007). Graphic organizers in reading instruction: Research findings and issues. *Reading in a Foreign Language*, 19(1), 34–55.
- Li, J. (2014). The role of reading and writing connections in ESL students' summarization strategies. *Reading and Writing*, 27(9), 1671–1689.
- Liu, J., & Read, J. (2020). Revisiting reading research for EFL/ESL tertiary learners: Where have we been and where are we going? *Language Teaching Research*, 26(5), 734–753.
- Lynn, C. A. (2010). Effects of graphic organizers on reading comprehension achievement of third grade students. (Doctoral dissertation, Liberty University).
- Mani, A., & Jammala, R. (2020). Challenges in English reading comprehension among Indian university students. *Asian EFL Journal*, 27(3), 88–109.
- Mede, E. (2010). The effects of graphic organizers on reading comprehension and attitudes of Turkish EFL learners. *Procedia – Social and Behavioral Sciences*, 2(2), 2705–2709.
- Qi, Y., & Jiang, Q. (2021). Application of graphic organizers in English reading teaching for students of various learning abilities. *Theory and Practice in Language Studies*, 11(9), 1097–1102.
- Sam, D. L., & Rajan, P. (2013). Using graphic organizers to improve reading comprehension skills for the middle school ESL students. *English Language Teaching*, 6(2), 155–169.
- Sanz-Tejeda, A., Domínguez-Oller, D., Baldaquí-Escandell, J., Gómez-Díaz, R., & García-Rodríguez, A. (2025). Artificial intelligence in reading comprehension and text production: A systematic review. *Computer Assisted Language Learning*. Advance online publication.
- Shafiee, S. (2025). AI-based technology and EFL reading comprehension: Recent trends and directions. *Language Teaching Research Quarterly*, 42, 78–95.

- Taylor, B. M. (1986). Teaching middle grade students to summarize content textbook material. In J. F. Baumann (Ed.), *Teaching main idea comprehension* (pp. 195–209). International Reading Association.
- Trehearne, M. P. (2006). *Comprehensive literacy resource for grades 3–6*. ETA Cuisenaire. (as cited in Bolambo, 2018).
- Warnby, M. (2025). *Reading for learning across the lifespan: Research perspectives*. Cambridge University Press.
- Zafarania, M., & Kabgani, S. (2014). The relationship between summarizing strategy use and reading comprehension ability. *Procedia – Social and Behavioral Sciences*, 98, 1959–1965.

Appendix 1: Analytical Scoring Rubric for Summarization Tasks

Criterion	Max Score	4–5 (Excellent)	2–3 (Adequate)	0–1 (Poor)	Score
Identification of main idea	5	Main idea clearly and accurately stated in own words	Main idea partially identified; some paraphrase	Main idea absent or copied verbatim	
Recognition of supporting details	5	Key details selected and integrated accurately	Some relevant details included; minor omissions	Details missing, irrelevant, or copied	
Vocabulary understanding	5	Content vocabulary used accurately and contextually	Vocabulary mostly accurate; some misuse	Vocabulary inaccurate or absent	
Inferencing	5	Implicit meanings drawn accurately and expressed clearly	Some inference evident; partially explained	No evidence of inferencing	
Total	20				

Note. Each passage was scored out of 5 points (reflecting a condensed application of the rubric), yielding a combined maximum of 15 points across all three passages. $\kappa = .81$ (inter-rater reliability, Cohen’s kappa, based on 20% stratified sample).