Voices of English Language Education Society
Vol. 10, No. 2 | August 2026 | E-ISSN 2579-7484
DOI: https://doi.org/10.29408/veles.v10i2.33627

Making Comparisons Visible: Effects of Venn Diagram Scaffolding on EFL Students' Critical Reading of Analytical Exposition Texts

Arinaa Sabilah Ghozali*[1], Siti Muniroh[1], Niamika El Khoiri[1]
[1]Malang State University, Indonesia
Received: 20 December 2025 | Revised: 27 May 2026 | Accepted: 05 August 2026
Correspondence: arinaa.sabilah.2302218@student.um.ac.id

Abstract

Critical reading requires EFL learners to evaluate and synthesize arguments while simultaneously processing an additional language. However, limited evidence exists regarding whether Venn Diagrams can scaffold the comparative analysis of analytical exposition texts. This study investigated the effect of a scaffolded Venn Diagram strategy on Indonesian EFL students' critical reading performance. A quasi-experimental pre-test–post-test design involved 60 Grade XI students from two intact classes at a private high school in Malang, Indonesia. The experimental group received five instructional sessions using Venn Diagrams, whereas the control group received conventional reading instruction. Critical reading was assessed using a ten-item essay test covering six cognitive levels and scored with an analytical rubric. The instrument demonstrated acceptable internal consistency (Cronbach's $\alpha = .761$), while inter-rater agreement exceeded $\kappa = .80$. Data were analyzed using descriptive statistics and independent-samples t-tests after normality and homogeneity assumptions were satisfied. The groups did not differ significantly at pre-test ($p = .981$). At post-test, however, the experimental group ($M = 88.43$, $SD = 8.01$) significantly outperformed the control group ($M = 73.20$, $SD = 10.44$), with a mean difference of 15.23 points, $t(58) = 6.342$, $p < .001$, Cohen's $d = 1.64$. The experimental group also achieved a larger mean gain (24.47 versus 9.87 points), accompanied by substantially reduced score variability. These findings indicate that scaffolded Venn Diagrams can externalize comparative relationships and support more effective and consistent analysis, synthesis, and evaluation of argumentative texts. The study extends graphic-organizer research beyond basic comparison and supports the integration of explicit visual scaffolding into EFL critical-reading instruction.

Keywords: analytical exposition; critical reading; EFL instruction; graphic organizer; quasi-experimental study; Venn Diagram

1. Introduction

Critical reading is a central component of academic literacy because learners must do more than retrieve explicitly stated information from a text; they must examine how meaning and arguments are constructed (Afflerbach et al., 2015; Sutherland & Incera, 2021; Van et al., 2022; Wilson, 2016). Effective critical readers distinguish claims from supporting evidence, compare competing ideas or viewpoints, identify unstated assumptions and logical inconsistencies, evaluate the relevance and credibility of evidence, and formulate conclusions grounded in textual information (Afflerbach et al., 2015; Din, 2020; Kiili et al., 2022; List & Alexander, 2019). These higher-order processes are particularly demanding for EFL learners, who must perform complex reasoning while simultaneously decoding unfamiliar vocabulary, processing syntactically complex sentences, and interpreting discourse relationships in an additional language (Siu & Ho, 2020; Wetzel et al., 2020; Yamashita & Shiotsu, 2017; Zhang & Zhang, 2022). The cognitive resources devoted to resolving these linguistic difficulties may limit learners' capacity to retain, connect, and critically assess multiple ideas across a text (Huang et al., 2022; List & Alexander, 2019; Shin, 2020; Shin et al., 2019). Consequently, students may understand individual sentences or locate factual information without recognizing weaknesses in an argument, differentiating evidence from opinion, or evaluating the writer's overall position (Afflerbach et al., 2015; Heidari, 2020; Kim, 2023; Le et al., 2024). This tension between linguistic comprehension and critical engagement highlights the need for reading instruction that supports learners not only in understanding what a text communicates but also in systematically examining, comparing, and judging the ideas it presents (Bakhtiari Moghadam et al., 2023; Hà & Huertas-Abril, 2024; Vongkulluksn et al., 2023; Wilson, 2016).

These challenges become particularly pronounced when learners read analytical exposition texts because comprehension requires them to construct the writer's argument rather than merely understand isolated propositions: readers must identify the thesis, trace relationships among claims, reasons, evidence, and examples, compare connected or competing ideas, and judge whether the conclusion follows from the preceding reasoning (Diakidoy et al., 2017; Kirpouiki & Agaliotis, 2025; Nussbaum et al., 2019; Tengberg & Olin-Scheller, 2016). For EFL learners, limitations in vocabulary, grammar, decoding, and word-to-text integration can widen the gap between sentence-level understanding and a coherent representation of the whole text; consequently, recognizing words, interpreting individual sentences, and locating explicit information may produce local understanding without ensuring comprehension of how evidence supports claims or how a conclusion reinforces the thesis (Lee et al., 2022; Li & Kirby, 2015; Mulder et al., 2021; Vettori et al., 2024). Such fragmented comprehension restricts learners' capacity to analyze, synthesize, and evaluate a text because higher-order reading depends on constructing and monitoring a coherent mental representation in which propositions are connected across sentences and sections (Clinton et al., 2020; Follmer, 2018; Helder et al., 2016; van den Broek & Helder, 2017). Across EFL settings, classroom and teacher-practice studies indicate that reading instruction may continue to foreground conventional skill work, vocabulary, and text-explicit comprehension while critical reading and higher-order analysis are integrated unevenly or incompletely; such lessons may check local understanding without explicitly teaching learners to map argumentative structure, compare claims, evaluate evidence, or formulate reasoned judgments (Al-Husban, 2019; Ansori et al., 2019; Gönen & Kızılay, 2022; Huh, 2016). Experimental and meta-analytic research demonstrates that explicit, systematic text-structure instruction can improve expository comprehension and higher-order strategy use, including among English learners (Bogaerds-Hazenberg et al., 2021; Hebert et al., 2016; Wijekumar et al., 2017; Xu et al., 2024). Taken together, research on argument mapping and visualization supports the use of an explicit scaffold that makes relationships among the thesis, supporting arguments, evidence, points of comparison or opposition, and conclusion visible, thereby helping learners progress from local comprehension toward an integrated critical evaluation of the argument (Chiang et al., 2016; Cullen et al., 2018; Eftekhari & Sotoudehnama, 2018; Mochizuki et al., 2019).

Addressing this difficulty calls for explicit instructional scaffolding because critical reading simultaneously draws on attention, information maintenance, and working-memory control; when task complexity exceeds available cognitive capacity, appropriately designed guidance can preserve resources for higher-order processing (Peng et al., 2018; Sweller et al., 2019; van Nooijen et al., 2024). To evaluate an analytical exposition text, readers must maintain the writer's thesis as an organizing representation, distinguish claims from premises and evidence, trace inferential relationships among supporting arguments, compare alternative positions, and determine whether the conclusion is warranted by the reasons advanced (Berkle et al., 2023; Christodoulou & Diakidoy, 2020; Diakidoy et al., 2015). For EFL learners, these operations can be further constrained by the cognitive effort required for lexical access, syntactic parsing, discourse-level integration, and the retention of information presented in a second language (Babayiğit & Shapiro, 2020; Kozan et al., 2015; Phillips Galloway & Uccelli, 2019). When these integrative processes are overtaxed and external support is unavailable, readers may fail to keep earlier propositions active, construct only local or fragmented text representations, and base their judgments on surface-level paraphrases or impressions rather than relationships across the argument (Latini et al., 2021; Martin et al., 2020; Tarchi et al., 2021). Instructional scaffolding can make this reasoning more manageable by decomposing critical reading into explicit, teachable operations—identifying claims, classifying and organizing evidence, comparing relevant features, and producing an evidence-based evaluative conclusion—while providing modeling and guided practice for each operation (Abrami et al., 2015; Du & List, 2022, 2024; van Peppen et al., 2018). Externalizing selected propositions and their relationships can further reduce the amount of information that learners must maintain internally, thereby freeing limited cognitive resources for comparison, inference, and judgment (Berry et al., 2019; Krieglstein et al., 2022; Risko & Gilbert, 2016). Such support should remain contingent and temporary: teachers initially model and prompt the target reasoning, guide learners as they apply it, and gradually fade assistance as learners assume greater responsibility, while

ensuring that support is withdrawn at an appropriate time (Ertugruloglu et al., 2023; Martin et al., 2019; van de Pol et al., 2015). Within this framework, graphic organizers are especially promising because they spatially display similarities, differences, hierarchies, and inferential connections, making otherwise abstract argumentative relationships more visible and available for critical examination (Dwyer et al., 2015; Kaeppel, 2021; Schroeder et al., 2018).

Among graphic organizers, Venn diagrams are especially well matched to critical-reading tasks that require systematic comparison because their intersecting regions make similarities, differences, and category membership explicit, although available evidence does not establish that they are consistently superior to other comparison organizers (Calvin & Gray, 2022; Katemba & Sihombing, 2023; Meyer et al., 2018; Pyle et al., 2017). When distinct claims, supporting evidence, or textual features are entered in the non-overlapping regions and shared elements are placed in the intersection, information dispersed across an analytical exposition can be reorganized into a coherent external representation, making patterns of convergence, divergence, and contradiction easier to inspect (Barzilai et al., 2018; Ciullo et al., 2015; Hattan et al., 2023; Urton et al., 2025). Because the diagram stores selected propositions and relationships externally and reduces repeated switching among separated textual elements, it may lower working-memory and attentional demands, leaving more cognitive capacity for analysis, synthesis, and evaluation (Gilbert et al., 2020; Morrison & Richmond, 2020; Paas & van Merriënboer, 2020; Pouw et al., 2019). These advantages are not automatic: research on learner-generated visual representations indicates that a blank or incomplete organizer can be ineffective and may create additional cognitive demands when students lack guidance concerning what information to select and how to organize it (Ainsworth & Scheiter, 2021; Colliot & Jamet, 2018; Colliot et al., 2022; Fiorella & Zhang, 2018). The intervention should therefore be operationalized as a scaffolded Venn Diagram strategy in which the teacher initially models how to establish relevant comparison criteria, distinguish claims from evidence, and determine where textual information belongs before providing supported practice in applying these procedures (Hall-Mills & Marante, 2022; Hall-Mills & Marante, 2025; Jian, 2021; Williams et al., 2016). Students can then be prompted to justify their placements, discuss relationships among ideas, and use the completed diagram as an evidentiary basis for evaluating the strength and coherence of the writer's argument, practices supported by research on self-explanation, deliberative dialogue, and source-based argument construction (Bisra et al., 2018; Casado-Ledesma et al., 2021; Rapanta & Felton, 2022; Shi & Iordanou, 2025). As competence develops, instructional support can be reduced and adjusted to learners' performance so that responsibility shifts toward independent construction and interpretation; nevertheless, because scaffolding effects vary with their design and context, fading should be treated as a monitored instructional decision rather than an automatic timetable (Belland et al., 2017; Doo et al., 2020; Kim et al., 2018; Zuo et al., 2023). Accordingly, the strategy's likely pedagogical value lies not merely in displaying similarities and differences but in combining visual externalization with explicit and dialogic reasoning that supports learners' progression from locating textual information to producing integrated, evidence-based evaluative conclusions (Cromley et al., 2020; Granado-Peinado et al., 2019; Mateos et al., 2018; Shi, 2019).

Research specifically examining Venn diagrams in EFL contexts remains limited and has focused predominantly on general comprehension rather than higher-order critical reading. In the study most closely aligned with the present research, Fauziah and Kumalarini (2017) observed the use of Venn diagrams with 46 Grade XI students reading analytical exposition texts and found that teacher modeling helped students compare textual ideas and that students perceived the strategy as facilitating comprehension; however, its descriptive design did not objectively test improvement or establish a treatment effect. More outcome-oriented studies have generally reported favorable results. Siregar (2018) found significantly higher reading-comprehension performance among eighth-grade students taught through Venn diagrams, while Rama (2018) similarly reported significantly greater gains for Indonesian junior-secondary learners receiving Venn-diagram instruction than for those receiving conventional instruction. At the senior-secondary level, Manurung et al. (2022) found that students taught using Venn diagrams significantly outperformed a conventionally instructed group on a multiple-choice reading test. Extending this evidence to an English-for-

specific-purposes context, Susanto (2021) reported larger comprehension gains among Indonesian law students using Venn diagrams to process legal descriptive texts. Venn diagrams have also supported idea generation and organization in EFL writing, as Syafii and Miftah (2020) found improved report-text writing and motivation across two classroom-action-research cycles. Nevertheless, the evidence is not uniformly conclusive: although the Venn-diagram group in Katemba and Sihombing's (2023) study improved from a pre-test mean of 71.22 to a post-test mean of 84.56, its gain did not differ significantly from that produced by KWL instruction. Moreover, a critical-reading intervention incorporating Venn diagrams alongside four other graphic organizers improved Saudi EFL learners' analysis, problem-solving, evaluation, and summarizing performance, but the bundled treatment prevented the specific contribution of the Venn diagram from being isolated (Barashid, 2020).

Previous research reveals substantive, methodological, and pedagogical gaps. Substantively, most EFL studies have examined Venn diagrams as tools for general reading comprehension, idea organization, or writing development, leaving their contribution to higher-order critical reading—particularly the analysis, synthesis, and evaluation of arguments—insufficiently understood. Methodologically, the study most closely related to analytical exposition employed a descriptive design, whereas other investigations frequently relied on multiple-choice comprehension measures or combined Venn diagrams with other graphic organizers, making it difficult to isolate their specific effect on critical-reading performance. Pedagogically, previous studies have generally treated the Venn diagram as a visual organizer rather than as part of an explicit reasoning process involving teacher modeling, guided comparison, justification of textual relationships, and gradual transfer of responsibility. To address these limitations, the present study investigated the effect of scaffolded Venn-diagram instruction on Indonesian secondary EFL students' critical reading of analytical exposition texts using a quasi-experimental pre-test–post-test design with a conventional-instruction comparison group. The study's novelty therefore lies not merely in introducing a familiar graphic organizer but in reconceptualizing it as a structured comparative-reasoning scaffold through which learners identify relevant arguments, distinguish similarities and differences, connect claims with textual evidence, synthesize dispersed information, and formulate evidence-based evaluations.

2. Method

2.1 Research Design

This study adopted a quantitative approach using a nonequivalent-groups pretest–posttest quasi-experimental design. This design was selected because the participants belonged to two intact Grade XI classes and could not be randomly assigned individually to the instructional conditions. Quasi-experimental designs are commonly used to evaluate educational interventions under authentic classroom conditions when random assignment is impractical, although researchers must account carefully for possible baseline differences and selection effects (Gopalan et al., 2020). One class was assigned as the experimental group and received scaffolded Venn-diagram instruction, while the other served as the control group and received conventional reading instruction. The instructional condition constituted the independent variable, whereas students' critical-reading performance constituted the dependent variable. Critical-reading performance was measured before and after the intervention using a constructed-response test. Table 1 summarizes the research design.

Table 1
Quasi-Experimental Research Design

Group	Pretest	Instructional condition	Posttest
Experimental	O_1	X_1: Scaffolded Venn-diagram instruction	O_2
Control	O_1	X_2: Conventional reading instruction	O_2

Note. O_1 = pretest; O_2 = posttest; X_1 = scaffolded Venn-diagram instruction; X_2 = conventional reading instruction.

2.2 Research Setting and Participants

The study was conducted at a private senior high school in Malang, Indonesia, whose curriculum emphasizes technology integration and literacy development. Data were collected from September to

October 2025. Each class participated in seven meetings: one pretest meeting, five instructional meetings, and one posttest meeting. Each meeting lasted 2 × 40 minutes in accordance with the school timetable. The participants were 60 Grade XI students from two intact classes, with 30 students in the experimental group and 30 in the control group. The classes were purposively selected based on the comparability of their previous English achievement scores and the compatibility of their schedules with the intervention. Grade XI was selected because analytical exposition texts formed part of the students' English curriculum and the instructional tasks were consistent with their prior English preparation. Because selection was conducted at the class rather than individual level, pretest scores were subsequently used to account for possible initial differences between the groups.

2.3 Research Instrument

Data were collected using a critical-reading test consisting of ten constructed-response questions. The test blueprint adopted the six cognitive categories applied in recent EFL critical-reading research: knowledge, comprehension, application, analysis, synthesis, and evaluation (Le et al., 2024). The items ranged from foundational comprehension tasks, such as determining phrase meaning and identifying the main idea, to higher-order tasks requiring students to distinguish facts from opinions, integrate ideas across texts, evaluate an author's position, and formulate evidence-based judgments. Constructed-response items were selected because students needed to articulate their interpretations and provide textual justification rather than recognize an answer from predetermined alternatives. Test format can influence the cognitive and linguistic demands of L2 reading assessment, making the alignment between response format and the targeted reading construct an important consideration (In'nami & Koizumi, 2009). Open-ended performance tasks are also appropriate when the targeted competencies involve interpretation, reasoning, and evaluation, provided that explicit scoring criteria are used to control scoring variability (Jonsson & Svingby, 2007). An English-education expert reviewed the test blueprint, reading passages, questions, answer guidelines, and scoring criteria. The review examined the representativeness of the items, their alignment with the targeted critical-reading indicators, clarity of wording, appropriateness for Grade XI students, and expected level of cognitive demand. This process provided evidence based on test content; expert judgment was treated as one component of the instrument's validity evidence rather than as definitive proof that the instrument was valid (Sireci, 1998). Table 2 presents the test blueprint.

Table 2
Blueprint of the Critical-Reading Test

Item	Cognitive category	Assessed task
1	Knowledge	Determine the meaning of a specified phrase.
2	Comprehension	Identify the main idea of the text.
3	Comprehension	Explain the communicative purpose of the text.
4	Analysis	Distinguish a factual statement from an opinion or evaluative claim.
5	Comprehension	Paraphrase a selected sentence to demonstrate understanding.
6	Evaluation	Judge the author's attitude and support the judgment with textual evidence.
7	Evaluation	Draw and justify a conclusion using relevant textual evidence.
8	Synthesis	Integrate ideas from two texts and explain their relationship.
9	Application	Apply arguments from the texts to predict possible real-world consequences.
10	Evaluation	Propose and justify appropriate actions based on principles presented in the texts.

Note. The instrument framework was adapted from Le et al. (2024).

Students' responses were evaluated using an analytical scoring rubric adapted from Elder and Paul's (2008) intellectual standards. The rubric assessed six dimensions of response quality: accuracy, clarity, precision, depth, relevance, and logic. The use of an analytical rubric enabled the raters to evaluate separate dimensions of performance instead of assigning a single holistic judgment. Research indicates that analytical and task-specific rubrics can improve the consistency of performance assessment when criteria are clearly defined and interpreted consistently by raters (Jonsson & Svingby, 2007). The maximum raw score was 68

points. Table 3 summarizes the scoring criteria and point allocation for each test item; the raw total was then converted to a percentage using the formula shown below.

Table 3
Analytical Scoring Rubric and Maximum Points per Item

Question	Rubric standard(s) and scoring focus	Maximum points
Q1	Clarity: Understanding the meaning of the specified phrase	4
Q2	Accuracy: Recognizing the main idea	4
Q3	Accuracy: Recognizing the author's purpose	4
Q4	Clarity: Identifying a factual statement	4
Q5	Precision: Rephrasing with appropriate vocabulary	4
Q6	Depth: Interpreting the author's attitude; Relevance: Providing a personal explanation or response	8 (4 + 4)
Q7	Depth: Connecting concepts from both texts; Relevance: Drawing a conclusion; Logic: Applying concepts	12 (4 + 4 + 4)
Q8	Depth: Connecting ideas across texts; Relevance: Synthesizing a conclusion; Logic: Applying concepts	12 (4 + 4 + 4)
Q9	Depth: Demonstrating a complex understanding of long-term impact; Logic: Applying concepts to a future scenario	8 (4 + 4)
Q10	Relevance: Generating ideas grounded in the text; Logic: Applying concepts to propose practical steps	8 (4 + 4)
	Total maximum points	**68**

$$\text{Final score (\%)} = (\text{Total points earned} / 68) \times 100$$

Score interpretation. A = Excellent (90–100%; 61–68 points); B = Good (80–89%; 54–60 points); C = Adequate (70–79%; 48–53 points); D = Needs Improvement (60–69%; 41–47 points); F = Insufficient (< 60%; 0–40 points).

The instrument was piloted with 30 students. Pearson item–total correlations were calculated to examine the relationship between each item score and the total test score. Using a critical value of $r = .361$, all ten items produced coefficients exceeding the threshold. These results were interpreted as evidence of satisfactory item functioning rather than as sufficient evidence of instrument validity, which must be supported by multiple sources of evidence (Sireci, 1998). Internal consistency was assessed using Cronbach's alpha, producing a coefficient of .761. This value was interpreted as evidence of adequate rather than "high" internal consistency because the interpretation of alpha should consider the number of items, construct breadth, and intended use of the scores (Taber, 2018). Interrater agreement was examined using Cohen's kappa, a chance-corrected agreement coefficient for categorical ratings (Hallgren, 2012). The coefficients exceeded .80 across the test administrations, indicating a high level of agreement between the two raters.

2.4 Instructional Procedure

The study was implemented in four stages. First, both groups completed the pretest under equivalent administration conditions. Second, the groups received five instructional meetings of equal duration. The experimental intervention followed the principles of scaffolding and gradual release of responsibility. Scaffolding involves providing support that is responsive to learners' needs, gradually reducing that support, and transferring responsibility for task performance to learners (van de Pol et al., 2010). The gradual-release framework similarly moves instruction from explicit teacher demonstration through supported participation toward increasingly independent performance, although its implementation should remain responsive rather than mechanically sequential (Webb et al., 2019).

In Meeting 1, the teacher introduced the communicative purpose, structure, and language features of analytical exposition texts. Meeting 2 involved explicit teacher modeling of how to construct and interpret a Venn diagram when comparing two analytical exposition texts. The teacher demonstrated how to establish relevant comparison criteria, identify the thesis and supporting arguments, distinguish claims from textual evidence, and place information in the appropriate sections of the diagram. Meeting 3 involved collaborative group analysis in which students constructed Venn diagrams with teacher guidance. Meeting 4 involved

paired analysis of more complex texts with reduced instructional support. In Meeting 5, students applied the strategy independently.

Throughout the intervention, students learned to compare the theses, supporting arguments, evidence, and relevant textual features presented in two analytical exposition texts. Information unique to each text was placed in the nonoverlapping sections of the diagram, while shared arguments or textual features were placed in the overlapping section. Students were required to justify each placement by referring to textual evidence and to use the completed diagram to formulate comparative and evaluative conclusions. Teacher support was gradually reduced as students progressed from modeled practice to collaborative, paired, and independent application.

During the same five-meeting period, the control group received conventional instruction in analytical exposition texts. The instruction included teacher explanations, whole-class discussions, vocabulary clarification, and standard comprehension questions without Venn diagrams or other graphic organizers. Both groups received texts of comparable length, linguistic difficulty, and thematic content and were allocated equal instructional time. Following the intervention, both groups completed the posttest under equivalent administration and scoring conditions. The responses were anonymized, scored by two raters, and prepared for statistical analysis.

2.5 Data Analysis

The data were analyzed using independent-samples t-tests to compare the experimental and control groups' pretest and posttest scores. The t-test was selected because the primary research question concerned whether there was a statistically significant difference between the two independent groups at each assessment point. Prior to conducting the t-tests, the distributional and variance assumptions were examined. Normality was evaluated using Q–Q plots and the Kolmogorov–Smirnov test, while homogeneity of variance was examined using Levene's test. Because normality tests differ in sensitivity across distributions and sample sizes, statistical results were considered together with graphical evidence rather than interpreted solely through a fixed sample-size rule (Yap & Sim, 2011). Descriptive statistics including means, standard deviations, minimum and maximum scores, and 95% confidence intervals were calculated for the pretest, posttest, and gain scores of both groups. Effect size was reported to complement statistical significance because a p value alone does not indicate the practical magnitude of an instructional effect. Cohen's d was calculated for the between-group difference and interpreted with reference to effect-size distributions commonly reported in second-language research (Plonsky & Oswald, 2014). All inferential tests were two-tailed and evaluated at $\alpha = .05$.

3. Findings

3.1 Descriptive Results and Assumption Testing

The experimental and control groups demonstrated nearly identical performance before the intervention. The experimental group obtained a pretest mean of 63.97 (SD = 21.87), while the control group obtained a mean of 63.83 (SD = 21.31). Following the five instructional meetings, the experimental group's mean increased to 88.43 (SD = 8.01), whereas the control group's mean increased to 73.20 (SD = 10.44). Thus, the raw mean increase was 24.47 points for the experimental group and 9.87 points for the control group. Table 4 summarizes the descriptive results.

Table 4

Descriptive Statistics for the Experimental and Control Groups

Group	Assessment	*n*	Minimum	Maximum	*M*	*SD*	Raw mean increase
Experimental	Pretest	30	20	93	63.97	21.87	—
Experimental	Posttest	30	67	100	88.43	8.01	24.47
Control	Pretest	30	20	93	63.83	21.31	—
Control	Posttest	30	55	93	73.20	10.44	9.87

The posttest scores of the experimental group were not only higher but also less dispersed than their pretest scores, as indicated by the decrease in the standard deviation from 21.87 to 8.01. The control group also showed a reduction in variability, although the change was smaller. These results indicate that the experimental group's posttest scores were more closely clustered around their mean. However, the reduction in standard deviation should not, by itself, be interpreted as evidence that the intervention equalized achievement across ability levels. Such a conclusion would require a direct analysis of treatment effects across initial-proficiency groups. The clustering may also partly reflect the experimental group's proximity to the upper limit of the 0–100 scale.

Before conducting the independent-samples *t*-tests, the distributional and variance assumptions were examined. The Kolmogorov–Smirnov tests did not indicate statistically significant departures from normality for either group at either assessment point. Levene's tests were also nonsignificant for the pretest and posttest scores, indicating that the equal-variance assumption was not violated. The results are presented in Table 5.

Table 5

Results of Normality and Homogeneity Tests

Assessment	Experimental K–S *p*	Control K–S *p*	Levene's *F*	Levene's *p*
Pretest	.142	≥ .200	0.026	.872
Posttest	≥ .200	≥ .200	3.254	.076

Note. K–S = Kolmogorov–Smirnov test. Values displayed by SPSS as .200 are reported as ≥ .200. All assumption-test values exceeded α = .05.

3.2 Between-Group Differences in Critical-Reading Performance

The independent-samples *t*-test found no statistically detectable pretest difference between the experimental and control groups, $t(58) = 0.024$, $p = .981$. The mean difference was only 0.14 points. Although a nonsignificant result does not formally prove that two groups are equivalent, the almost identical means and standard deviations provide no indication of a meaningful baseline imbalance in the reported scores.

A different pattern emerged at posttest. The experimental group's mean score of 88.43 was 15.23 points higher than the control group's mean of 73.20. This difference was statistically significant, $t(58) = 6.342$, $p < .001$, with a 95% confidence interval of approximately [10.42, 20.04]. The standardized difference was large, Cohen's $d = 1.64$. Accordingly, the null hypothesis of no posttest difference between the groups was rejected. Table 6 presents the complete comparison.

Table 6

Independent-Samples t-Test Results

Assessment	Experimental *M* (*SD*)	Control *M* (*SD*)	Mean difference	95% CI	*t*	*df*	*p*	Cohen's *d*
Pretest	63.97 (21.87)	63.83 (21.31)	0.14	[−11.02, 11.30]	0.024	58	.981	0.01
Posttest	88.43 (8.01)	73.20 (10.44)	15.23	[10.42, 20.04]	6.342	58	< .001	1.64

Note. Mean differences were calculated as experimental minus control. Cohen's *d* was calculated using the pooled standard deviation.

The results show that students who received scaffolded Venn-diagram instruction completed the intervention with substantially higher critical-reading test scores than students who received conventional reading instruction. Both groups demonstrated descriptive improvement, suggesting that both instructional conditions supported learning to some extent. Nevertheless, the experimental group's raw mean increase was approximately 15 points greater than that of the control group, and the large standardized posttest difference indicates that the observed advantage was educationally meaningful as well as statistically significant.

4. Discussion

The present study found that students receiving scaffolded Venn-diagram instruction achieved substantially stronger critical-reading performance than those receiving conventional instruction. Before the intervention, the experimental and control groups had nearly identical mean scores of 63.97 and 63.83, respectively. After five instructional meetings, the experimental group obtained a mean of 88.43, compared

with 73.20 for the control group. The 15.23-point posttest difference was statistically significant, $t(58) = 6.342$, $p < .001$, and corresponded to a large standardized difference, Cohen's $d = 1.64$. Both groups improved descriptively, but the experimental group's raw increase of 24.47 points was considerably greater than the control group's increase of 9.87 points. Together with the close baseline scores, these results provide evidence that the scaffolded Venn-diagram condition offered an instructional advantage over the conventional approach used in this study. This finding is broadly consistent with research showing that visual comparison organizers can support reading comprehension. Calvin and Gray (2022), for example, found that explicit instruction in double-bubble maps improved bilingual middle-school students' comprehension of compare-and-contrast texts. Katemba and Sihombing (2023) similarly reported improvement following Venn-diagram instruction among Indonesian secondary-school students. However, they found no significant difference between Venn-diagram and KWL instruction, indicating that Venn diagrams may not necessarily outperform another structured and active reading strategy. More broadly, meta-analytic evidence indicates that studying and constructing relational visual maps can produce meaningful learning benefits by making connections among ideas more explicit (Schroeder et al., 2018). The present findings therefore support Venn diagrams as an effective structured alternative to the conventional instruction employed here, rather than establishing their universal superiority over all other reading strategies.

One explanation for the experimental group's stronger performance is the alignment between the spatial structure of the Venn diagram and the comparative reasoning required by the critical-reading tasks. Students had to identify the thesis, supporting arguments, and evidence in two analytical exposition texts, distinguish information unique to each text, and determine where the texts converged. The diagram converted these relationships into a visible structure: separate sections represented distinctive claims or evidence, while the intersection represented shared positions or argumentative features. Research consistently demonstrates that explicit attention to expository-text structure can improve readers' identification, organization, and comprehension of important information (Bogaerds-Hazenberg et al., 2021; Hall-Mills & Marante, 2022; Hebert et al., 2016; Pyle et al., 2017). Similar benefits have been reported when text-structure strategy instruction is applied to English-language learners, suggesting that explicit organizational support can help L2 readers construct more coherent representations of expository texts (Xu et al., 2024).

The visual representation itself may also have helped students coordinate information from the texts with the relationships displayed in the diagram. Learning from visual representations requires readers to select relevant information, interpret the diagram's spatial organization, and integrate the visual and verbal representations into a coherent mental model (Ainsworth & Scheiter, 2021; Hattan et al., 2023; Jian, 2021). In the present intervention, these processes were supported through teacher modeling and guided practice rather than left entirely to students. Consequently, the Venn diagram did not merely summarize information after reading; it provided an organizing structure through which students could progressively construct their interpretation of the two texts. Nevertheless, organizing information is not sufficient by itself to constitute critical reading. Critical reading requires readers to move beyond literal comprehension by examining how claims are supported, evaluating the relevance and credibility of evidence, identifying assumptions or inconsistencies, and forming justified conclusions (Afflerbach et al., 2015; Hà & Huertas-Abril, 2024; Sutherland & Incera, 2021; Van et al., 2022). This distinction is particularly important for argumentative texts because accurately recalling a writer's position does not automatically enable a reader to evaluate the quality of the argument. Argument knowledge and an explicit evaluative orientation are important for recognizing argumentative structures, judging the strength of reasons, and detecting weaknesses or fallacies (Berkle et al., 2023; Christodoulou & Diakidoy, 2020; Diakidoy et al., 2017).

The requirement that students justify where they placed information may therefore have been a crucial component of the intervention. Rather than simply listing similarities and differences, students had to explain how particular claims, reasons, or pieces of evidence belonged in the overlapping or nonoverlapping sections. This process made relationships among claims, reasons, and evidence available for examination and

discussion. Studies of argument visualization have found that visually mapping argumentative relationships can support argumentative-text comprehension, analytical reasoning, recall, and reflective judgment (Chiang et al., 2016; Cullen et al., 2018; Dwyer et al., 2015; Eftekhari & Sotoudehnama, 2018). Although argument maps are structurally more elaborate than Venn diagrams, these findings support the broader explanation that externalizing relationships among argumentative elements can facilitate the analysis of complex texts. The use of two texts also positioned the intervention as a form of multiple-text reading. Reading multiple texts requires more than understanding each text independently; readers must determine how their claims relate, identify agreement and contradiction, and construct an integrated representation across sources. Reviews of multiple-text comprehension have identified comparison, cross-textual organization, and integration as essential but demanding processes (Barzilai et al., 2018; List & Alexander, 2019; Tarchi et al., 2021). Instruction that explicitly directs learners to integrate conflicting information, connect claims with evidence, and construct cross-text conclusions has been shown to strengthen argumentative synthesis and critical integrative reasoning (Du & List, 2022; Mateos et al., 2018; Vongkulluksn et al., 2023). The Venn diagram's overlapping and nonoverlapping sections may have supported these processes by giving students a stable location in which to represent convergence and divergence before formulating their conclusions.

A further plausible explanation concerns the processing demands of L2 critical reading. Students had to decode English texts while retaining claims, comparing details, evaluating evidence, and constructing written responses. Working-memory capacity is associated with reading comprehension generally and with inferential comprehension among L2 readers specifically (Kim, 2023; Shin, 2020; Vettori et al., 2024). By recording relevant information externally, the diagram may have reduced the need to maintain all claims and supporting details simultaneously in working memory. This interpretation is consistent with cognitive-offloading research, which explains how restructuring the external environment can reduce internal memory demands, and with cognitive-load theory, which emphasizes the value of instructional representations that direct limited cognitive resources toward task-relevant processing (Berry et al., 2019; Paas & van Merriënboer, 2020; Risko & Gilbert, 2016; Sweller et al., 2019). However, cognitive load and working memory were not measured in this study. Reduced processing demand should therefore be treated as a theoretically plausible mechanism rather than a demonstrated explanation for the score difference. The findings should also be attributed to the complete scaffolded intervention rather than to the Venn diagram in isolation. The experimental condition combined the visual organizer with explicit teacher modeling, collaborative analysis, paired practice, evidence-based justification, and gradual movement toward independent application. Research indicates that critical-thinking and critical-reading outcomes are more likely to improve when reasoning strategies are taught explicitly, demonstrated in context, practiced repeatedly, and accompanied by opportunities for guided application (Abrami et al., 2015; Bakhtiari Moghadam et al., 2023; Tengberg & Olin-Scheller, 2016; Wilson, 2016). Research on scaffolding further emphasizes that instructional materials are most effective when teachers provide responsive support and progressively adjust that support as learners become more capable (Ertugruloglu et al., 2023; Martin et al., 2019; van de Pol et al., 2015). The observed effect consequently represents the value of a scaffolded Venn-diagram strategy, not merely the distribution of blank diagrams to students.

This distinction may also help explain why the experimental group's standard deviation decreased from 21.87 at pretest to 8.01 at posttest. The smaller posttest standard deviation indicates that students' scores became more closely concentrated around the higher mean. It does not, however, demonstrate that the intervention functioned as a "great equalizer" or benefited students at every initial-proficiency level equally. The clustering may partly reflect the group's proximity to the upper limit of the test, and an ability-by-treatment analysis would be required to establish whether initially weaker students benefited more strongly. Research also shows that the effectiveness of graphic organizers can depend on their completeness, complexity, and the amount of guidance provided during construction (Colliot & Jamet, 2018; Colliot et al., 2022; Fiorella & Zhang, 2018; Krieglstein et al., 2022). Thus, the reduction in variability supports the consistency of the observed posttest performance but should not be interpreted as direct evidence of an equalizing effect. Pedagogically, the findings suggest that Venn diagrams should be used as reasoning tools

rather than as worksheets for listing surface-level similarities and differences. Teachers should explicitly model how to establish meaningful comparison criteria, identify theses and supporting arguments, distinguish claims from evidence, and justify why information belongs in a particular section. Students should then use the completed diagram to formulate comparative conclusions, evaluate the relative strength of arguments, and defend their judgments with textual evidence. Requiring explanations is important because self-explanation and critical questioning can promote deeper processing, transfer, and evidence-based argument evaluation (Bisra et al., 2018; Nussbaum et al., 2019; Shi, 2019; van Peppen et al., 2018). The gradual progression from modeled to collaborative, paired, and independent analysis may consequently help students internalize a repeatable procedure for approaching unfamiliar argumentative texts.

Several limitations qualify these conclusions. The study involved only two intact classes from one school, individual random assignment was not conducted, and the intervention lasted five instructional meetings. The aggregate test score also does not demonstrate that every critical-reading component—such as identifying claims, integrating evidence, recognizing bias, and evaluating arguments—improved to the same degree. Item-level or dimension-level analyses would be required to determine which aspects of critical reading changed most strongly. Moreover, the study did not directly measure strategy use, cognitive load, or the quality of students' completed diagrams, so the mechanisms proposed in this discussion remain inferential. Future research should incorporate adjusted posttest analysis, delayed assessment, process measures, and comparisons with other structured strategies such as KWL or argument mapping. Within these limitations, the findings provide encouraging evidence that scaffolded Venn-diagram instruction can support EFL secondary students' critical reading of analytical exposition texts. Its apparent value lies not simply in presenting information visually but in combining comparative representation with explicit instruction, evidence-based justification, and progressively independent analysis. The study therefore supports the Venn diagram as part of an integrated critical-reading pedagogy that makes relationships among claims, arguments, and evidence more visible and manageable for EFL learners.

5. Conclusion

This study examined the effectiveness of scaffolded Venn-diagram instruction in developing EFL secondary students' critical reading of analytical exposition texts. After five instructional meetings, the experimental group improved from 63.97 to 88.43, whereas the control group improved from 63.83 to 73.20. The significant posttest difference, $t(58) = 6.342$, $p < .001$, with a large effect size ($d = 1.64$), indicates that the intervention was more effective than the conventional instruction used in this study. The findings suggest that the value of the Venn diagram lies not in the visual template alone but in its scaffolded use through teacher modeling, guided practice, textual justification, and gradual movement toward independent analysis. This approach helped students organize similarities and differences across texts and examine relationships among claims, evidence, and perspectives more systematically. Nevertheless, the use of two intact classes from one school and the short intervention period limit the generalizability of the findings. Future research should employ larger, multisite samples, delayed assessments, process-based measures, and comparisons with other active reading strategies. Within these limitations, the study provides encouraging evidence that scaffolded Venn-diagram instruction can strengthen critical reading by making comparative and evaluative reasoning more visible and manageable for EFL learners.

References

Abrami, P. C., Bernard, R. M., Borokhovski, E., Waddington, D. I., Wade, C. A., & Persson, T. (2015). Strategies for teaching students to think critically: A meta-analysis. *Review of Educational Research, 85*(2), 275–314. https://doi.org/10.3102/0034654314551063

Afflerbach, P., Cho, B.-Y., & Kim, J.-Y. (2015). Conceptualizing and assessing higher-order thinking in reading. *Theory Into Practice, 54*(3), 203–212. https://doi.org/10.1080/00405841.2015.1044367

Ainsworth, S. E., & Scheiter, K. (2021). Learning by drawing visual representations: Potential, purposes, and practical implications. *Current Directions in Psychological Science, 30*(1), 61–67. https://doi.org/10.1177/0963721420979582

Al-Husban, N. (2019). EFL teachers' practices while teaching reading comprehension in Jordan: Teacher development implications. *Journal on English as a Foreign Language, 9*(2), 127–145. https://doi.org/10.23971/jefl.v9i2.1288

Ansori, M., Nurkamto, J., & Suparno, S. (2019). Teacher's beliefs and practices in the integration of higher order thinking skills in teaching reading. *ELS Journal on Interdisciplinary Studies in Humanities, 2*(4), 541–555. https://doi.org/10.34050/els-jish.v2i4.8164

Babayiğit, S., & Shapiro, L. R. (2020). Component skills that underpin listening comprehension and reading comprehension in learners with English as first and additional language. *Journal of Research in Reading, 43*(1), 78–97. https://doi.org/10.1111/1467-9817.12291

Bakhtiari Moghadam, Z., Narafshan, M. H., & Tajadini, M. (2023). The effect of implementing a critical thinking intervention program on English language learners' critical thinking, reading comprehension, and classroom climate. *Asian-Pacific Journal of Second and Foreign Language Education, 8*, Article 15. https://doi.org/10.1186/s40862-023-00188-3

Barashid, M. (2020). Investigating the critical reading skills for Saudi EFL learners through graphic organizers instructional program. *Frontiers in Education Technology, 3*(4), 30–42. https://doi.org/10.22158/fet.v3n4p30

Barzilai, S., Zohar, A. R., & Mor-Hagani, S. (2018). Promoting integration of multiple texts: A review of instructional approaches and practices. *Educational Psychology Review, 30*(3), 973–999. https://doi.org/10.1007/s10648-018-9436-8

Belland, B. R., Walker, A. E., Kim, N. J., & Lefler, M. (2017). Synthesizing results from empirical research on computer-based scaffolding in STEM education: A meta-analysis. *Review of Educational Research, 87*(2), 309–344. https://doi.org/10.3102/0034654316670999

Berkle, Y., Schmitt, L., Tolzin, A., Janson, A., Wambsganss, T., Leimeister, J. M., & Leuchter, M. (2023). Measuring university students' ability to recognize argument structures and fallacies. *Frontiers in Psychology, 14*, Article 1270931. https://doi.org/10.3389/fpsyg.2023.1270931

Berry, E. D. J., Allen, R. J., Mon-Williams, M., & Waterman, A. H. (2019). Cognitive offloading: Structuring the environment to improve children's working memory task performance. *Cognitive Science, 43*(8), Article e12770. https://doi.org/10.1111/cogs.12770

Bisra, K., Liu, Q., Nesbit, J. C., Salimi, F., & Winne, P. H. (2018). Inducing self-explanation: A meta-analysis. *Educational Psychology Review, 30*(3), 703–725. https://doi.org/10.1007/s10648-018-9434-x

Bogaerds-Hazenberg, S. T. M., Evers-Vermeul, J., & van den Bergh, H. (2021). A meta-analysis on the effects of text structure instruction on reading comprehension in the upper elementary grades. *Reading Research Quarterly, 56*(3), 435–462. https://doi.org/10.1002/rrq.311

Calvin, K. L., & Gray, S. (2022). Double-bubble Thinking Maps and their effect on reading comprehension in Spanish–English bilingual middle school students with learning disabilities. *Learning Disability Quarterly, 45*(3), 212–224. https://doi.org/10.1177/0731948720958644

Casado-Ledesma, L., Cuevas, I., Van den Bergh, H., Rijlaarsdam, G., Mateos, M., Granado-Peinado, M., & Martín, E. (2021). Teaching argumentative synthesis writing through deliberative dialogues: Instructional practices in secondary education. *Instructional Science, 49*(4), 515–559. https://doi.org/10.1007/s11251-021-09548-3

Chiang, K.-H., Fan, C.-Y., Liu, H.-H., & Chen, G.-D. (2016). Effects of a computer-assisted argument map learning strategy on sixth-grade students' argumentative essay reading comprehension. *Multimedia Tools and Applications, 75*(16), 9973–9990. https://doi.org/10.1007/s11042-015-2904-y

Christodoulou, S. A., & Diakidoy, I.-A. N. (2020). The contribution of argument knowledge to the comprehension and critical evaluation of argumentative text. *Contemporary Educational Psychology, 63*, Article 101903. https://doi.org/10.1016/j.cedpsych.2020.101903

Ciullo, S., Falcomata, T. S., Pfannenstiel, K., & Billingsley, G. (2015). Improving learning with science and social studies text using computer-based concept maps for students with disabilities. *Behavior Modification, 39*(1), 117–135. https://doi.org/10.1177/0145445514552890

Clinton, V., Taylor, T., Bajpayee, S., Davison, M. L., Carlson, S. E., & Seipel, B. (2020). Inferential comprehension differences between narrative and expository texts: A systematic review and meta-analysis. *Reading and Writing, 33*(9), 2223–2248. https://doi.org/10.1007/s11145-020-10044-2

Colliot, T., & Jamet, É. (2018). Does self-generating a graphic organizer while reading improve students' learning? *Computers & Education, 126*, 13–22. https://doi.org/10.1016/j.compedu.2018.06.028

Colliot, T., Kiewra, K. A., Luo, L., Flanigan, A. E., Lu, J., Kennedy, C., & Black, S. (2022). The effects of graphic organizer completeness and note-taking medium on computer-based learning. *Education and Information Technologies, 27*(2), 2435–2456. https://doi.org/10.1007/s10639-021-10693-y

Cromley, J. G., Du, Y., & Dane, A. P. (2020). Drawing-to-learn: Does meta-analysis show differences between technology-based drawing and paper-and-pencil drawing? *Journal of Science Education and Technology, 29*(2), 216–229. https://doi.org/10.1007/s10956-019-09807-6

Cullen, S., Fan, J., van der Brugge, E., & Elga, A. (2018). Improving analytical reasoning and argument understanding: A quasi-experimental field study of argument visualization. *npj Science of Learning, 3*, Article 21. https://doi.org/10.1038/s41539-018-0038-5

Diakidoy, I.-A. N., Christodoulou, S. A., Floros, G., Iordanou, K., & Kargopoulos, P. V. (2015). Forming a belief: The contribution of comprehension to the evaluation and persuasive impact of argumentative text. *British Journal of Educational Psychology, 85*(3), 300–315. https://doi.org/10.1111/bjep.12074

Diakidoy, I.-A. N., Ioannou, M. C., & Christodoulou, S. A. (2017). Reading argumentative texts: Comprehension and evaluation goals and outcomes. *Reading and Writing, 30*(9), 1869–1890. https://doi.org/10.1007/s11145-017-9757-x

Din, M. (2020). Evaluating university students' critical thinking ability as reflected in their critical reading skill: A study at bachelor level in Pakistan. *Thinking Skills and Creativity, 35*, Article 100627. https://doi.org/10.1016/j.tsc.2020.100627

Doo, M. Y., Bonk, C. J., & Heo, H. (2020). A meta-analysis of scaffolding effects in online learning in higher education. *The International Review of Research in Open and Distributed Learning, 21*(3), 60–80. https://doi.org/10.19173/irrodl.v21i3.4638

Du, H., & List, A. (2022). Reasoning about text-based evidence. *Contemporary Educational Psychology, 68*, Article 102038. https://doi.org/10.1016/j.cedpsych.2021.102038

Du, H., & List, A. (2024). Evidence-based reasoning: Results from an intervention. *Applied Cognitive Psychology, 38*(5), Article e4238. https://doi.org/10.1002/acp.4238

Dwyer, C. P., Hogan, M. J., & Stewart, I. (2015). The effects of argument mapping-infused critical thinking instruction on reflective judgement performance. *Thinking Skills and Creativity, 16*, 11–26. https://doi.org/10.1016/j.tsc.2014.12.002

Eftekhari, M., & Sotoudehnama, E. (2018). Effectiveness of computer-assisted argument mapping for comprehension, recall, and retention. *ReCALL, 30*(3), 337–354. https://doi.org/10.1017/S0958344017000337

Elder, L., & Paul, R. (2008). *The thinker's guide to intellectual standards: The words that name them and the criteria that define them*. Foundation for Critical Thinking Press.

Ertugruloglu, E., Mearns, T., & Admiraal, W. (2023). Scaffolding what, why and how? A critical thematic review study of descriptions, goals, and means of language scaffolding in bilingual education contexts. *Educational Research Review, 40*, Article 100550. https://doi.org/10.1016/j.edurev.2023.100550

Fauziah, D. A., & Kumalarini, T. (2017). The implementation of Venn diagram as a strategy in teaching reading. *RETAIN, 5*(3), 226–233.

Fiorella, L., & Zhang, Q. (2018). Drawing boundary conditions for learning by drawing. *Educational Psychology Review, 30*(3), 1115–1137. https://doi.org/10.1007/s10648-018-9444-8

Follmer, D. J. (2018). Executive function and reading comprehension: A meta-analytic review. *Educational Psychologist, 53*(1), 42–60. https://doi.org/10.1080/00461520.2017.1309295

Gilbert, S. J., Bird, A., Carpenter, J. M., Fleming, S. M., Sachdeva, C., & Tsai, P.-C. (2020). Optimal use of reminders: Metacognition, effort, and cognitive offloading. *Journal of Experimental Psychology: General, 149*(3), 501–517. https://doi.org/10.1037/xge0000652

Gönen, S. İ. K., & Kızılay, Y. (2022). Reading beyond the lines: Teaching critical reading in higher education. *The Reading Matrix: An International Online Journal, 22*(1), 73–92. Full text

Gopalan, M., Rosinger, K., & Ahn, J. B. (2020). Use of quasi-experimental research designs in education research: Growth, promise, and challenges. *Review of Research in Education, 44*(1), 218–243. https://doi.org/10.3102/0091732X20903302

Granado-Peinado, M., Mateos, M., Martín, E., & Cuevas, I. (2019). Teaching to write collaborative argumentative syntheses in higher education. *Reading and Writing, 32*(8), 2037–2058. https://doi.org/10.1007/s11145-019-09939-6

Hà, T. A., & Huertas-Abril, C. A. (2024). Teaching critical reading in higher education: A literature review and pedagogical framework proposal. *Thinking Skills and Creativity, 53*, Article 101599. https://doi.org/10.1016/j.tsc.2024.101599

Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: An overview and tutorial. *Tutorials in Quantitative Methods for Psychology, 8*(1), 23–34. https://doi.org/10.20982/tqmp.08.1.p023

Hall-Mills, S. S., & Marante, L. M. (2022). Explicit text structure instruction supports expository text comprehension for adolescents with learning disabilities: A systematic review. *Learning Disability Quarterly, 45*(1), 55–68. https://doi.org/10.1177/0731948720906490

Hall-Mills, S. S., & Marante, L. M. (2025). Teaching expository text management and proficiency skills for comprehension for students with language/learning disabilities. *Learning Disability Quarterly, 48*(2), 88–101. https://doi.org/10.1177/07319487221145689

Hattan, C., Lee, E., & List, A. (2023). Comprehension, diagram analysis, integration, and interest: A cross-sectional analysis. *Reading Psychology, 44*(7), 731–760. https://doi.org/10.1080/02702711.2023.2187907

Hebert, M., Bohaty, J. J., Nelson, J. R., & Brown, J. (2016). The effects of text structure instruction on expository reading comprehension: A meta-analysis. *Journal of Educational Psychology, 108*(5), 609–629. https://doi.org/10.1037/edu0000082

Heidari, K. (2020). Critical thinking and EFL learners' performance on textually explicit, textually implicit, and script-based reading items. *Thinking Skills and Creativity, 37*, Article 100703. https://doi.org/10.1016/j.tsc.2020.100703

Helder, A., Van Leijenhorst, L., & van den Broek, P. (2016). Coherence monitoring by good and poor comprehenders in elementary school: Comparing offline and online measures. *Learning and Individual Differences, 48*, 17–23. https://doi.org/10.1016/j.lindif.2016.02.008

Huang, L., Ouyang, J., & Jiang, J. (2022). The relationship of word processing with L2 reading comprehension and working memory: Insights from eye-tracking. *Learning and Individual Differences, 95*, Article 102143. https://doi.org/10.1016/j.lindif.2022.102143

Huh, S. (2016). Instructional model of critical literacy in an EFL context: Balancing conventional and critical literacy. *Critical Inquiry in Language Studies, 13*(3), 210–235. https://doi.org/10.1080/15427587.2016.1154445

In'nami, Y., & Koizumi, R. (2009). A meta-analysis of test format effects on reading and listening test performance: Focus on multiple-choice and open-ended formats. *Language Testing, 26*(2), 219–244. https://doi.org/10.1177/0265532208101006

Jian, Y.-C. (2021). The immediate and delayed effects of text–diagram reading instruction on reading comprehension and learning processes: Evidence from eye movements. *Reading and Writing, 34*(3), 727–752. https://doi.org/10.1007/s11145-020-10089-3

Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review, 2*(2), 130–144. https://doi.org/10.1016/j.edurev.2007.05.002

Kaeppel, K. (2021). The influence of collaborative argument mapping on college students' critical thinking about contentious arguments. *Thinking Skills and Creativity, 40*, Article 100809. https://doi.org/10.1016/j.tsc.2021.100809

Katemba, C. V., & Sihombing, N. M. (2023). Enhancing students' reading comprehension through Venn Diagram strategy and Know, Want, Learn (KWL) strategy. *Acuity: Journal of English Language Pedagogy, Literature and Culture, 8*(1), 87–100. https://doi.org/10.35974/acuity.v8i1.3015

Kiili, C., Bråten, I., Strømsø, H. I., Hagerman, M. S., Räikkönen, E., & Jyrkiäinen, A. (2022). Adolescents' credibility justifications when evaluating online texts. *Education and Information Technologies, 27*(6), 7421–7450. https://doi.org/10.1007/s10639-022-10907-x

Kim, M. (2023). Exploring literal and inferential reading comprehension among L2 adolescent learners: The roles of working memory capacity, syllogistic inference, and L2 linguistic knowledge. *Reading and Writing, 36*(5), 1085–1110. https://doi.org/10.1007/s11145-022-10320-3

Kim, N. J., Belland, B. R., & Walker, A. E. (2018). Effectiveness of computer-based scaffolding in the context of problem-based learning for STEM education: Bayesian meta-analysis. *Educational Psychology Review, 30*(2), 397–429. https://doi.org/10.1007/s10648-017-9419-1

Kirpouiki, A., & Agaliotis, I. (2025). Teaching argumentative text comprehension to secondary students with specific reading disability or low reading achievement. *British Journal of Special Education, 52*(1), 81–90. https://doi.org/10.1111/1467-8578.12575

Kisbu-Sakarya, Y., MacKinnon, D. P., & Aiken, L. S. (2013). A Monte Carlo comparison study of the power of the analysis of covariance, simple difference, and residual change scores in testing two-wave data. *Educational and Psychological Measurement, 73*(1), 47–62. https://doi.org/10.1177/0013164412450574

Kozan, K., Erçetin, G., & Richardson, J. C. (2015). Input modality and working memory: Effects on second language text comprehension in a multimedia learning environment. *System, 55*, 63–73. https://doi.org/10.1016/j.system.2015.09.001

Krieglstein, F., Schneider, S., Beege, M., & Rey, G. D. (2022). How the design and complexity of concept maps influence cognitive learning processes. *Educational Technology Research and Development, 70*(1), 99–118. https://doi.org/10.1007/s11423-022-10083-2

Latini, N., Bråten, I., & Haverkamp, Y. E. (2021). Breadth and depth of strategic processing during text comprehension. *Learning and Individual Differences, 91*, Article 102058. https://doi.org/10.1016/j.lindif.2021.102058

Le, H. V., Nguyen, T. A. D., Le, D. H. N., Nguyen, P. U., & Nguyen, T. T. A. (2024). Unveiling critical reading strategies and challenges: A mixed-methods study among English major students in a Vietnamese higher education institution. *Cogent Education, 11*(1), Article 2326732. https://doi.org/10.1080/2331186X.2024.2326732

Lee, H., Jung, G., & Lee, J. H. (2022). Simple view of second language reading: A meta-analytic structural equation modeling approach. *Scientific Studies of Reading, 26*(6), 585–603. https://doi.org/10.1080/10888438.2022.2087526

Li, M., & Kirby, J. R. (2015). The effects of vocabulary breadth and depth on English reading. *Applied Linguistics, 36*(5), 611–634. https://doi.org/10.1093/applin/amu007

List, A., & Alexander, P. A. (2019). Toward an integrated framework of multiple text use. *Educational Psychologist, 54*(1), 20–39. https://doi.org/10.1080/00461520.2018.1505514

Manurung, L. W., Saragih, F., Saragi, C. N., & Manurung, L. (2022). The effect of applying Venn diagram strategy on students' achievement in reading comprehension at the twelfth grade students of SMK N 5 Medan. *Jurnal Scientia, 11*(1), 537–544. https://doi.org/10.35337/scientia.Vol11.pp537-544

Martin, J. D., Shipstead, Z., Harrison, T. L., Redick, T. S., Bunting, M., & Engle, R. W. (2020). The role of maintenance and disengagement in predicting reading comprehension and vocabulary learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 46*(1), 140–154. https://doi.org/10.1037/xlm0000705

Martin, N. D., Dornfeld Tissenbaum, C., Gnesdilow, D., & Puntambekar, S. (2019). Fading distributed scaffolds: The importance of complementarity between teacher and material scaffolds. *Instructional Science, 47*(1), 69–98. https://doi.org/10.1007/s11251-018-9474-0

Mateos, M., Martín, E., Cuevas, I., Villalón, R., Martínez, I., & González-Lamas, J. (2018). Improving written argumentative synthesis by teaching the integration of conflicting information from multiple sources. *Cognition and Instruction, 36*(2), 119–138. https://doi.org/10.1080/07370008.2018.1425300

Meyer, B. J. F., Wijekumar, K., & Lei, P.-W. (2018). Comparative signaling generated for expository texts by 4th–8th graders: Variations by text structure strategy instruction, comprehension skill, and signal word. *Reading and Writing, 31*(9), 1937–1968. https://doi.org/10.1007/s11145-018-9871-4

Mochizuki, T., Nishimori, T., Tsubakimoto, M., Oura, H., Sato, T., Johansson, H., Nakahara, J., & Yamauchi, Y. (2019). Development of software to support argumentative reading and writing by means of creating a graphic organizer from an electronic text. *Educational Technology Research and Development, 67*(5), 1197–1230. https://doi.org/10.1007/s11423-019-09676-1

Morrison, A. B., & Richmond, L. L. (2020). Offloading items from memory: Individual differences in cognitive offloading in a short-term memory task. *Cognitive Research: Principles and Implications, 5*(1), Article 1. https://doi.org/10.1186/s41235-019-0201-4

Mulder, E., van de Ven, M., Segers, E., Krepel, A., de Bree, E. H., de Jong, P. F., & Verhoeven, L. (2021). Word-to-text integration in English as a second language reading comprehension. *Reading and Writing, 34*(4), 1049–1087. https://doi.org/10.1007/s11145-020-10097-3

Nussbaum, E. M., Dove, I. J., Slife, N., Kardash, C. M., Turgut, R., & Vallett, D. (2019). Using critical questions to evaluate written and oral arguments in an undergraduate general education seminar: A quasi-experimental study. *Reading and Writing, 32*(6), 1531–1552. https://doi.org/10.1007/s11145-018-9848-3

Paas, F., & van Merriënboer, J. J. G. (2020). Cognitive-load theory: Methods to manage working-memory load in the learning of complex tasks. *Current Directions in Psychological Science, 29*(4), 394–398. https://doi.org/10.1177/0963721420922183

Peng, P., Barnes, M., Wang, C., Wang, W., Li, S., Swanson, H. L., Dardick, W., & Tao, S. (2018). A meta-analysis on the relation between reading and working memory. *Psychological Bulletin, 144*(1), 48–76. https://doi.org/10.1037/bul0000124

Phillips Galloway, E., & Uccelli, P. (2019). Examining developmental relations between core academic language skills and reading comprehension for English learners and their peers. *Journal of Educational Psychology, 111*(1), 15–31. https://doi.org/10.1037/edu0000276

Plonsky, L., & Oswald, F. L. (2014). How big is "big"? Interpreting effect sizes in L2 research. *Language Learning, 64*(4), 878–912. https://doi.org/10.1111/lang.12079

Pouw, W. T. L. J., Rop, G., de Koning, B. B., & Paas, F. (2019). The cognitive basis for the split-attention effect. *Journal of Experimental Psychology: General, 148*(11), 2058–2075. https://doi.org/10.1037/xge0000578

Pyle, N., Vasquez, A. C., Lignugaris/Kraft, B., Gillam, S. L., Reutzel, D. R., Olszewski, A., Segura, H., Hartzheim, D., Laing, W., & Pyle, D. (2017). Effects of expository text structure interventions on comprehension: A meta-analysis. *Reading Research Quarterly, 52*(4), 469–501. https://doi.org/10.1002/rrq.179

Rama, A. N. (2018). The effect of Venn diagram to enhance students reading comprehension at the second year of SMP Negeri 1 Wawotobi. *Cetta: Jurnal Ilmu Pendidikan, 1*(2), 64–75.

Rapanta, C., & Felton, M. K. (2022). Learning to argue through dialogue: A review of instructional approaches. *Educational Psychology Review, 34*(2), 477–509. https://doi.org/10.1007/s10648-021-09637-2

Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. *Trends in Cognitive Sciences, 20*(9), 676–688. https://doi.org/10.1016/j.tics.2016.07.002

Schroeder, N. L., Nesbit, J. C., Anguiano, C. J., & Adesope, O. O. (2018). Studying and constructing concept maps: A meta-analysis. *Educational Psychology Review, 30*(2), 431–455. https://doi.org/10.1007/s10648-017-9403-9

Shi, Y. (2019). Enhancing evidence-based argumentation in a Mainland China middle school. *Contemporary Educational Psychology, 59*, Article 101809. https://doi.org/10.1016/j.cedpsych.2019.101809

Shi, Y., & Iordanou, K. (2025). A dialogic approach to enhancing argumentative writing from multiple texts. *ECNU Review of Education, 8*(2), 530–556. https://doi.org/10.1177/20965311241310884

Shin, J. (2020). A meta-analysis of the relationship between working memory and second-language reading comprehension: Does task type matter? *Applied Psycholinguistics, 41*(4), 873–900. https://doi.org/10.1017/S0142716420000272

Shin, J., Dronjic, V., & Park, B. (2019). The interplay between working memory and background knowledge in L2 reading comprehension. *TESOL Quarterly, 53*(2), 320–347. https://doi.org/10.1002/tesq.482

Sireci, S. G. (1998). The construct of content validity. *Social Indicators Research, 45*, 83–117. https://doi.org/10.1023/A:1006985528729

Siregar, S. R. (2018). The effect of Venn diagram strategy to students reading comprehension ability at eight grade of SMP Swasta Nurul Ilmi Padangsidimpuan. *English Education: English Journal for Teaching and Learning, 6*(1), 39–51. https://doi.org/10.24952/ee.v6i1.1217

Siu, T.-S. C., & Ho, S.-H. C. (2020). A longitudinal investigation of syntactic awareness and reading comprehension in Chinese–English bilingual children. *Learning and Instruction, 67*, Article 101327. https://doi.org/10.1016/j.learninstruc.2020.101327

Susanto, F. (2021). The effect of "The Venn Diagram" as a cooperative learning strategy to improve reading comprehension in legal description texts during the COVID-19 pandemic period: As a simple quasi-experimental approach with a non-equivalent control group in English specific purpose design for the Sariah Faculty of Constitutional Law. *Edu-Ling: Journal of English Education and Linguistics, 5*(1), 80–90. https://doi.org/10.32663/edu-ling.v5i1.2524

Sutherland, A., & Incera, S. (2021). Critical reading: What do faculty think students should do? *Journal of College Reading and Learning, 51*(4), 267–290. https://doi.org/10.1080/10790195.2021.1887777

Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review, 31*(2), 261–292. https://doi.org/10.1007/s10648-019-09465-5

Syafii, M. L., & Miftah, M. Z. (2020). Venn-diagram strategy in EFL class to enhance learners' writing skill and motivation. *Journal on English as a Foreign Language, 10*(1), 145–167. https://doi.org/10.23971/jefl.v10i1.1556

Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education, 48*(6), 1273–1296. https://doi.org/10.1007/s11165-016-9602-2

Tarchi, C., Ruffini, C., & Pecini, C. (2021). The contribution of executive functions when reading multiple texts: A systematic literature review. *Frontiers in Psychology, 12*, Article 716463. https://doi.org/10.3389/fpsyg.2021.716463

Tengberg, M., & Olin-Scheller, C. (2016). Developing critical reading of argumentative text: Effects of a comprehension strategy intervention. *Journal of Language Teaching and Research, 7*(4), 635–645. https://doi.org/10.17507/jltr.0704.02

Urton, K., Moeyaert, M., Nobel, K., Barwasser, A., Boon, R. T., & Grünke, M. (2025). Effects of graphic organizers on outcomes for students with disabilities: Three-level meta-analysis of single-case studies. *Exceptionality, 33*(1), 17–39. https://doi.org/10.1080/09362835.2024.2389080

van de Pol, J., Volman, M., & Beishuizen, J. (2010). Scaffolding in teacher–student interaction: A decade of research. *Educational Psychology Review, 22*(3), 271–296. https://doi.org/10.1007/s10648-010-9127-6

van de Pol, J., Volman, M., Oort, F., & Beishuizen, J. (2015). The effects of scaffolding in the classroom: Support contingency and student independent working time in relation to student achievement, task effort and appreciation of support. *Instructional Science, 43*(5), 615–641. https://doi.org/10.1007/s11251-015-9351-z

van den Broek, P., & Helder, A. (2017). Cognitive processes in discourse comprehension: Passive processes, reader-initiated processes, and evolving mental representations. *Discourse Processes, 54*(5–6), 360–372. https://doi.org/10.1080/0163853X.2017.1306677

van Nooijen, C. C. A., de Koning, B. B., Bramer, W. M., Isahakyan, A., Asoodar, M., Kok, E., van Merriënboer, J. J. G., & Paas, F. (2024). A cognitive load theory approach to understanding expert scaffolding of visual problem-solving tasks: A scoping review. *Educational Psychology Review, 36*, Article 12. https://doi.org/10.1007/s10648-024-09848-3

van Peppen, L. M., Verkoeijen, P. P. J. L., Heijltjes, A. E. G., Janssen, E. M., Koopmans, D., & van Gog, T. (2018). Effects of self-explaining on learning and transfer of critical thinking skills. *Frontiers in Education, 3*, Article 100. https://doi.org/10.3389/feduc.2018.00100

Van, L. H., Li, C. S., & Wan, R. (2022). Critical reading in higher education: A systematic review. *Thinking Skills and Creativity, 44*, Article 101028. https://doi.org/10.1016/j.tsc.2022.101028

Vettori, G., Casado Ledesma, L., Tesone, S., & Tarchi, C. (2024). Key language, cognitive and higher-order skills for L2 reading comprehension of expository texts in English as foreign language students: A systematic review. *Reading and Writing, 37*(9), 2481–2519. https://doi.org/10.1007/s11145-023-10479-3

Vongkulluksn, V. W., Nishiyama, C., Rodriguez, M. C., & Nussbaum, E. M. (2023). Critical Reading of Informational Texts (CRIT) scaffold: Evaluating the efficacy of an instructional scaffold for reading multiple scientific texts. *Contemporary Educational Psychology, 75*, Article 102229. https://doi.org/10.1016/j.cedpsych.2023.102229

Webb, S., Massey, D., Goggans, M., & Flajole, K. (2019). Thirty-five years of the gradual release of responsibility: Scaffolding toward complex and responsive teaching. *The Reading Teacher, 73*(1), 75–83. https://doi.org/10.1002/trtr.1799

Wetzel, M., Zufferey, S., & Gygax, P. (2020). Second-language acquisition and the mastery of discourse connectives: Assessing the factors that hinder L2 learners from mastering French connectives. *Languages, 5*(3), Article 35. https://doi.org/10.3390/languages5030035

Wijekumar, K. K., Meyer, B. J. F., & Lei, P. (2017). Web-based text structure strategy instruction improves seventh graders' content area reading comprehension. *Journal of Educational Psychology, 109*(6), 741–760. https://doi.org/10.1037/edu0000168

Williams, J. P., Kao, J. C., Pao, L. S., Ordynans, J. G., Atkins, J. G., Cheng, R., & DeBonis, D. (2016). Close analysis of texts with structure (CATS): An intervention to teach reading comprehension to at-risk second graders. *Journal of Educational Psychology, 108*(8), 1061–1077. https://doi.org/10.1037/edu0000117

Wilson, K. (2016). Critical reading, critical thinking: Delicate scaffolding in English for Academic Purposes (EAP). *Thinking Skills and Creativity, 22*, 256–265. https://doi.org/10.1016/j.tsc.2016.10.002

Xu, Z., Wijekumar, K. K., Wang, Q., Irey, R., & Liang, H. (2024). The effects of web-based text structure strategy instruction on adult Chinese ELLs' reading comprehension and reading strategy use. *Language Teaching Research, 28*(4), 1288–1310. https://doi.org/10.1177/13621688211022308

Yamashita, J., & Shiotsu, T. (2017). Comprehension and knowledge components that predict L2 reading: A latent-trait approach. *Applied Linguistics, 38*(1), 43–67. https://doi.org/10.1093/applin/amu079

Yap, B. W., & Sim, C. H. (2011). Comparisons of various types of normality tests. *Journal of Statistical Computation and Simulation, 81*(12), 2141–2155. https://doi.org/10.1080/00949655.2010.520163

Zhang, S., & Zhang, X. (2022). The relationship between vocabulary knowledge and L2 reading/listening comprehension: A meta-analysis. *Language Teaching Research, 26*(4), 696–725. https://doi.org/10.1177/1362168820913998

Zuo, M., Kong, S., Ma, Y., Hu, Y., & Xiao, M. (2023). The effects of using scaffolding in online learning: A meta-analysis. *Education Sciences, 13*(7), Article 705. https://doi.org/10.3390/educsci13070705