Jupiter: Jurnal Pendidikan Terapan
Homepage : https://journal.diginus.id/jupiter/index
Email: jurnalpendidikanterapan@gmail.com

p-ISSN : 2985-3214; e-ISSN : 2964-3171
Jupiter, Vol. 4, No. 3, July 2026
Page 566-573

Diagnostic Profile of Inferential Reading Ability across Four Indicators among Fifth Graders in Bandung Elementary Schools

Indra Suhendra[1*], Wahyu Sopandi[2], Isah Cahyani[3], Cepi Riyana[4]
[1,2,3,4]Universitas Pendidikan Indonesia, Bandung, Indonesia

ARTICLE INFO

Article history:
Submitted: Juny 20, 2026
Final Revised: July 03, 2026
Accepted: July 11, 2026
Published: July 31, 2026

Keywords:
Inferential Reading; Diagnostic Profile; Reading Comprehension; Elementary Students.

ABSTRACT

Purpose: This study aimed to construct a diagnostic profile of inferential reading ability across four indicators anaphoric resolution, background knowledge activation, predictive reasoning, and retrospective synthesis among fifth-grade students in Bandung elementary schools, addressing the limited availability of dimension-specific empirical data on inferential reading performance at the primary level. **Methods**: A descriptive survey was conducted with 84 fifth-grade students from three public elementary schools in Bandung, who completed an inferential reading test structured around the four indicators. A broader cohort of 185 students from the same schools responded to a perception questionnaire, while 35 teachers completed a questionnaire on instructional practices. Semi-structured interviews were also conducted with three classroom teachers and three students. Descriptive statistics, frequency distributions, cross-school comparisons, and thematic analysis were employed. **Findings**: The overall mean score was 45.83 (SD = 15.94), with the majority of students performing below expected levels. Retrospective synthesis recorded the lowest mean (M = 39.06, 84.5% in the lowest category), followed by predictive reasoning and background knowledge activation (78.6% each). Anaphoric resolution, although the highest (M = 54.69), still left more than half of the students below the expected threshold. Comparable patterns were observed across the three schools. **Research Implications**: The findings indicate that inferential reading competence requires more systematic instructional attention, particularly in retrospective synthesis and predictive reasoning. The four-indicator diagnostic framework offers educators a tool for pinpointing dimension-specific weaknesses and provides researchers a baseline for future intervention studies. **Originality**: This study contributes an empirically grounded diagnostic profile of inferential reading ability across four discrete indicators at the elementary level, integrating objective test data with teacher and student perceptual evidence within a single analytical framework.

Doi: https://doi.org/10.61255/jupiter.v4i3.1411

INTRODUCTION

Reading comprehension is widely recognized as a foundational skill that determines academic success across all subject areas in elementary education (Dixon & Oakhill, 2024; Oakhill et al., 2014; Oakhill & Cain, 2019). A critical dimension of reading comprehension is inferential ability, defined as the capacity to construct meaning beyond explicitly stated information by connecting textual elements with each other and with the reader's prior knowledge (Kintsch, 1998; Snow, 2002). Research has demonstrated that children who struggle with inferential comprehension often perform adequately on literal tasks, leading teachers to overestimate their actual reading competence (Oakhill & Cain, 2019). This phenomenon, sometimes described as "word-level decoders, text-level non-comprehenders," has been documented across multiple languages and educational systems (Hulme & Snowling, 2011; Nation et al., 2010). Without diagnostic instruments that specifically target inferential dimensions, reading difficulties may go undetected until students face increasingly complex academic texts in higher grades, at which point remediation becomes substantially more challenging.

Researchers have identified several distinct cognitive operations that constitute inferential reading. Four indicators have received particular attention in the assessment literature: anaphoric resolution, the ability to resolve pronoun and lexical references across sentences (Lata et al., 2021; Sanford & Garrod, 1989), background knowledge activation, which involves retrieving and applying relevant prior knowledge to fill conceptual gaps in the text (Elbro & Buch-Iversen, 2013), predictive reasoning, defined as anticipating upcoming textual content based on partial cues and logical progression (Attaprechakul, 2013), and retrospective synthesis, the capacity to integrate information from earlier and later sections of a text into a coherent overall understanding (Fox & Alexander, 2014). Although analytically distinct, these four indicators are rarely examined simultaneously within a single diagnostic framework, particularly in non-Western educational settings.

Previous research on inferential reading in the Indonesian context has predominantly employed intervention-based designs, focusing on the effectiveness of instructional strategies such as reciprocal teaching and collaborative strategic reading (Rahmawati et al., 2021; Wahyudi & Rozimela, 2024). While these studies have contributed valuable pedagogical insights, they share a common limitation: they measure inferential reading improvement as a unitary construct without disaggregating performance across specific inferential operations. Consequently, the literature provides little guidance on which inferential dimensions pose the greatest difficulty for elementary students. Only a limited number of studies have attempted to establish a diagnostic baseline of inferential reading ability at the elementary level (Ardiansyah & Hardhienata, 2025). This gap is consequential because without a clear empirical profile, intervention designs risk being generic rather than targeted, and educators lack a diagnostic vocabulary for discussing students' specific reading difficulties.

The present study was designed to address this gap by constructing a diagnostic profile of inferential reading ability across four indicators among fifth-grade students in three public elementary schools in Bandung, Indonesia. Fifth grade represents a critical transition point at which students are expected to shift from learning to read toward reading to learn (Chall, 1983; Palumbo & Sanacore, 2009). In the Indonesian curriculum, fifth-grade students are expected to identify main ideas, draw conclusions, and relate textual information to personal experience, skills that presuppose competent inferential processing (Juliati et al., 2025). The study was guided by the following research questions:

1. What is the diagnostic profile of fifth-grade students' inferential reading ability across four indicators, namely anaphoric resolution, background knowledge activation, predictive reasoning, and retrospective synthesis?
2. How do the ability patterns across these four indicators compare among the three participating schools?
3. How do teacher and student perceptions corroborate the test-based diagnostic profile?

To answer these questions, the study employed a multi-instrument approach combining an objective inferential reading test with teacher and student questionnaires and semi-structured interviews. This design allows the findings to serve as both a diagnostic reference for practitioners and an empirical foundation for designing targeted interventions in subsequent research.

METHOD

This study employed a descriptive design with multi-source data collection to construct a diagnostic profile of inferential reading ability among fifth-grade students. The design was selected because the research objective was not to test causal relationships but to document and map the current state of inferential reading performance across four specific indicators, thereby establishing an empirical baseline for future experimental work (Creswell, 2002; Creswell & Creswell, 2022). Although the study incorporated quantitative data from tests and questionnaires alongside qualitative data from interviews, it remained a descriptive inquiry rather than a fully integrated mixed-methods study. The quantitative and qualitative data were analyzed separately and subsequently compared to identify areas of convergence and divergence between measured performance and perceptual accounts.

The study was conducted in three public elementary schools located in Zona D, Kota Bandung, selected through purposive sampling. These schools were chosen based on two criteria: they had not previously implemented any specialized inferential reading program, and their principals expressed willingness to participate in a diagnostic assessment of reading competence. The primary sample consisted of 84 fifth-grade students drawn evenly across the three schools (School A, n = 29; School B, n = 28; School C, n = 27). Additionally, 35 classroom teachers from these schools completed a teacher questionnaire on instructional practices and perceived reading difficulties. The student questionnaire was administered to a broader cohort of 185 fifth-grade students from the same three schools, which included the 84 test participants. Semi-structured interviews were conducted with three teachers and three students, purposively selected to represent varied levels of teaching experience and self-reported reading ability.

The primary instrument was an inferential reading test designed to measure student performance across four indicators: anaphoric resolution, background knowledge activation, predictive reasoning, and retrospective synthesis.

The test comprised four short narrative and expository passages, each followed by indicator-specific items, yielding a total of 20 items with a maximum score of 100. Each indicator was represented by five items worth five points each. Items required students to perform the target cognitive operation: resolving a pronoun or lexical reference across sentences for anaphoric resolution, retrieving and applying prior knowledge for background knowledge activation, anticipating subsequent events based on textual clues for predictive reasoning, and integrating information across earlier and later text segments for retrospective synthesis. Each item was scored using a rubric ranging from 0 (no response or completely incorrect) to 5 (fully correct and well-articulated), with intermediate scores for partially correct responses. The test was developed by the researchers and reviewed by two experts in reading assessment and elementary education. The Content Validity Index (CVI) was 0.91, and item-total correlations exceeded 0.30 for all items. The test was piloted with 30 fifth-grade students from a school outside the main sample. Cronbach's alpha for the full test was 0.82, indicating acceptable internal consistency, with subscale reliability values ranging from 0.71 to 0.78.

Supporting instruments included a teacher questionnaire measuring three dimensions: perceived student difficulty in inferential reading (8 items), perceived student independent learning (6 items), and attitudes toward technology in reading instruction (5 items), all rated on a five-point Likert scale. The student questionnaire covered parallel dimensions plus demographic items (20 items total). Semi-structured interview protocols for teachers explored instructional strategies, perceived barriers to inferential reading development, and assessment practices, while student protocols addressed reading habits, difficulties with specific text types, and self-perceived competence. All instruments were piloted with the same group of 30 students and five teachers outside the main sample and refined prior to full administration.

Data collection proceeded over a period of four weeks. Ethical approval was obtained from the institutional research ethics committee of Universitas Pendidikan Indonesia. Written informed consent was secured from school principals, participating teachers, and parents or guardians of all participating students. Students provided verbal assent prior to each data collection activity. In the first week, the teacher and student questionnaires were distributed and collected. In the second week, the inferential reading test was administered under standardized conditions during regular class hours, with a duration of approximately 60 minutes per class. In the third and fourth weeks, semi-structured interviews were conducted individually in a quiet room within each school, lasting between 20 and 30 minutes per participant. All interviews were audio-recorded with participant consent and transcribed verbatim.

For the inferential reading test, descriptive statistics were computed for each indicator and for the total score, including means, standard deviations, and frequency distributions. Student performance was classified into four criterion-based categories aligned with conventional grading standards in Indonesian elementary education: Poor (0 to 59), Fair (60 to 74), Good (75 to 89), and Excellent (90 to 100). These boundaries provide a meaningful interpretation of performance relative to the test scale rather than relative to the sample distribution. Cross-school comparisons were presented through descriptive tables to examine whether performance patterns were consistent across school contexts. Questionnaire data were analyzed using descriptive statistics, with teacher and student responses categorized into Low, Medium, and High levels based on mean scores and standard deviation ranges. Interview transcripts were analyzed thematically following the six-phase framework of Braun and Clarke (2006). To enhance credibility, two researchers independently coded 20% of the transcripts, yielding an intercoder agreement of 88%, with discrepancies resolved through discussion. Findings from the interviews were triangulated with the quantitative results from the test and questionnaires.

RESULTS

The inferential reading test was administered to 84 fifth-grade students across three elementary schools. Table 1 presents the descriptive statistics for each indicator.

Table 1. Descriptive Statistics of Inferential Reading Test Scores (N = 84)

Indicator	Min	Max	Mean	SD
Anaphoric	0	100.00	54.69	18.16
Background Knowledge	0	100.00	46.88	19.78
Predictive	0	75.00	42.71	19.14
Retrospective	0	100.00	39.06	20.82

The overall mean score across all indicators was 45.83 (SD = 15.94), indicating that students, on average, performed below the midpoint of the scale. Retrospective recorded the lowest mean (M = 39.06, SD = 20.82), while Anaphoric recorded the highest (M = 54.69, SD = 18.16). Predictive recorded a maximum score of only 75.00, the lowest ceiling among all indicators. The standard deviations were relatively large across all indicators, suggesting considerable

variation within the sample. To provide a more interpretable picture of student performance, scores were classified into four criterion-based categories using the following boundaries: Poor (0 to 59), Fair (60 to 74), Good (75 to 89), and Excellent (90 to 100). Table 2 displays the distribution of students across these categories for each indicator.

Table 2. Distribution of Student Categories per Indicator (N = 84)

Indicator	Poor (0–59) %	Fair (60–74) %	Good (75–89) %	Excellent (90–100) %
Anaphoric	58.3	32.1	6.0	3.6
Background Knowledge	78.6	11.9	6.0	3.6
Predictive	78.6	20.2	1.2	0.0
Retrospective	84.5	11.9	1.2	2.4

On the total score, 77.4% of students fell into the Poor category, and no student reached the Excellent threshold. The indicator with the highest concentration of Poor performance was Retrospective, where 84.5% of students were classified at the lowest level. Predictive recorded no students at the Excellent level. Even Anaphoric, the strongest indicator, left 58.3% of students in the Poor category. Only on Anaphoric and Background Knowledge did any students reach Excellent, each with three students (3.6%). A cross-school comparison was conducted to examine whether the performance pattern varied across schools. Table 3 presents the mean total score per school.

Table 3. Mean Total Score per School

School	N	Mean	SD
School A	29	44.36	16.81
School B	28	48.03	13.63
School C	27	45.15	18.05

School B recorded the highest mean (M = 48.03, SD = 13.63), while School A recorded the lowest (M = 44.36, SD = 16.81). The spread across schools was narrow, with all three means falling within a range of less than four points and none reaching the Fair category threshold. Similar patterns were observed across the three schools.

Teacher questionnaire data were consistent with the test findings. Among the 35 participating teachers, responses to the Inferential Reading dimension (8 items, five-point scale) recorded a mean of 3.93 (SD = 0.68). Using the categorization described in the Method section, 22.9% of teachers rated student competence as Low and 62.9% as Medium. Teachers identified tasks requiring students to connect disparate parts of a text (M = 3.88) and find implied meanings (M = 3.86) as the areas of greatest student difficulty.

Student questionnaire data, collected from 185 respondents across the same three schools, revealed a parallel pattern. The Inferential Reading dimension (5 items, five-point scale) registered a mean of 2.91 (SD = 0.89), with all five items clustering between 2.90 and 2.94. Only 18.4% of students fell into the High category for self-assessed inferential ability, whereas 58.4% fell into the High category for enthusiasm toward technology use in learning.

Semi-structured interviews with three teachers yielded six themes: students could read aloud fluently but struggled to explain implied meanings; teacher perceptions of student understanding often overestimated actual comprehension; students waited for teacher instructions rather than initiating reading tasks independently; no systematic strategy had been employed to teach inferential skills; teachers were enthusiastic about digital tools but lacked ready-made resources; and similar patterns of difficulty were reported across all three schools.

Interviews with three students yielded five themes: reading words was easy but understanding implied meaning was difficult; they used spontaneous strategies such as skipping unknown words rather than deliberate comprehension monitoring; they rarely evaluated their own understanding after finishing a text; they primarily used mobile devices for gaming and social media rather than learning; and they expressed strong interest in a digital learning assistant that could respond to their questions at any time.

DISCUSSION

The present study set out to construct a diagnostic profile of inferential reading ability across four indicators among fifth-grade students in three Bandung elementary schools. The findings reveal a consistent pattern: inferential reading competence in this sample fell substantially below expected levels across all measured dimensions, and the difficulty was distributed in a patterned rather than random fashion.

The four-indicator profile shows a clear gradient of difficulty. Anaphoric resolution, which requires tracking pronoun references and lexical substitutions across adjacent sentences, produced the highest mean score and the lowest proportion of students in the Poor category. This outcome is consistent with developmental models of inference-making, which position anaphoric processing as the earliest-acquired and most text-proximal inferential operation

(Eilers et al., 2019). Resolving a pronoun reference demands that the reader hold two pieces of explicitly available information simultaneously, the pronoun and its antecedent, without requiring prior knowledge activation or global text integration (Badecker & Straub, 2002; Patterson, 2013). Even this relatively undemanding operation, however, left nearly six in ten students classified as Poor, indicating that inferential difficulties in this sample extended even to text-proximal operations.

Background knowledge activation and predictive reasoning occupied the middle tier of the profile, with identical proportions of Poor-level performance but meaningfully different means (Corkery, 2014; Hughson, 2003). The finding that most students could not adequately activate prior knowledge to support text comprehension is consistent with research showing that many elementary-level readers fail to deploy relevant background knowledge spontaneously during reading, even when that knowledge is accessible to them (Elbro & Buch-Iversen, 2013). Predictive reasoning posed a somewhat different challenge. Generating a prediction requires constructing a mental model of the text up to a given point, identifying its logical trajectory, and projecting that trajectory forward. The absence of any student reaching the Excellent category on this indicator suggests that predictive engagement during reading does not develop incidentally through the classroom practices experienced by these students. The lower ceiling observed on Predictive (maximum score of 75.00) compared to other indicators also raises the possibility that the predictive reasoning items were inherently more demanding than items in other indicator sets, a question that warrants further investigation through formal item analysis.

Retrospective synthesis emerged as the most severe point of difficulty, with a mean score of 39.06 and 84.5% of students in the Poor category. This indicator demands the highest level of integrative processing: the reader must retain information from earlier text segments, connect it with later segments, and construct a coherent representation of the whole (van den Broek et al., 2015). Unlike anaphoric resolution, which operates on adjacent sentences, retrospective synthesis spans the entire text. Unlike background knowledge activation, which draws on resources external to the text, retrospective synthesis requires strictly text-internal integration. The finding that this indicator was the weakest of the four aligns with theoretical accounts that characterize global coherence-building as the most cognitively demanding component of reading comprehension (Trabasso & Magliano, 1996). Taken together, the four indicators suggest a broader pattern in which difficulty increases as the inferential operation moves further from the surface of the text, a trend consistent with findings from reading comprehension research across different age groups and languages (Cain & Oakhill, 2007). This interpretation, however, should be approached with caution, as differences in item difficulty across the four indicator sets may have contributed to the observed score gradient and cannot be ruled out without formal item-level analysis.

The triangulation of test data, questionnaire data, and interview data strengthens the diagnostic picture. Teachers across all three schools rated inferential reading difficulty as the highest concern among the measured dimensions, a perception broadly aligned with the test results. Students, however, rated their inferential difficulties at a moderate level, a self-assessment that is markedly inconsistent with the objective finding that most of them performed at the Poor level. This gap between perceived and actual competence mirrors a well-documented pattern in reading research: younger readers, particularly those with reading difficulties, tend to overestimate their comprehension (Cain & Oakhill, 2007). The interview data offer a possible explanation: students reported that they rarely evaluated their own understanding after finishing a text. If reaching the end of a passage is equated with understanding it, the monitoring processes that would otherwise signal a comprehension breakdown may not be activated. It must be emphasized, however, that metacognitive monitoring was not directly measured in this study, and the link between self-assessment accuracy and monitoring behavior remains a hypothesis for future research.

The qualitative findings from teacher interviews add further context. Teachers described a consistent pattern in which students could read aloud fluently but could not explain what the text meant beyond its literal content, an observation that closely mirrors the "word-level decoders, text-level non-comprehenders" phenomenon documented in multiple educational settings (Hulme & Snowling, 2011; Nation et al., 2010). Teachers also reported that no systematic strategy for teaching inferential skills had been implemented in any of the three schools. While this finding offers a plausible instructional explanation for the test results, it is important to note that the present study employed a descriptive design; causal links between instructional practices and student performance cannot be established from these data.

Several limitations should be considered when interpreting these findings. The sample, although drawn from three different schools, was confined to a single geographic zone within one city and selected through purposive sampling based on principals' willingness to participate. This approach may have introduced selection bias, and the results should not be generalized to elementary school populations in rural or more remote regions where access to reading materials, teacher qualifications, and home-language dynamics may differ substantially. The study also did not measure variables that are known to influence inferential reading performance, including students' general language

ability, family literacy environment, and exposure to diverse text types. Classroom-level differences in reading instruction were not systematically documented, which limits the capacity to link specific teaching practices to the observed performance patterns. The inferential reading test, although reviewed by experts and piloted with a small group, was developed specifically for this study and has not undergone large-scale psychometric validation. Each indicator was measured by only five items, and text-level factors such as passage length, genre, and linguistic complexity were not experimentally controlled. Variation in item difficulty across indicators may therefore have contributed to the observed performance profile. Future instrument refinement should incorporate formal item analysis, a larger calibration sample, and systematic examination of text-level variables. Finally, the cross-sectional, descriptive design captures a single point in time and does not permit causal inference or developmental analysis. Longitudinal and experimental studies are needed to determine whether the four-indicator profile is stable over time and whether targeted instruction can alter the gradient of difficulty observed here.

Notwithstanding these limitations, the study makes several contributions. It provides a dimension-specific empirical baseline of inferential reading ability at the elementary level, addressing a gap that has been noted but infrequently examined empirically (Widiastuti, 2025; Zhou et al., 2025). The four-indicator framework offers a diagnostic vocabulary that enables educators to move beyond the broad label of "reading difficulty" toward identifying which particular inferential operations require targeted attention. The triangulation of test, questionnaire, and interview data yields a multi-layered profile that captures both measured performance and the perceptions of teachers and students, an approach that enriches the diagnostic picture beyond what test scores alone can provide.

The implications for practice should be advanced cautiously, given the study's descriptive scope and limited sample. Teachers in settings comparable to the participating schools may benefit from incorporating instructional strategies that explicitly target the inferential operations identified as most challenging, particularly retrospective synthesis and predictive reasoning. Activities such as constructing graphic organizers that map relationships across extended passages and engaging students in structured prediction tasks with justification may help address these specific weaknesses. Equally important is the cultivation of comprehension monitoring, a skill that the interview data suggest was largely absent among the participating students. Future research employing intervention designs will be necessary to evaluate whether these instructional approaches can produce measurable improvements in the inferential dimensions profiled in this study.

CONCLUSION

This study constructed a diagnostic profile of inferential reading ability across four indicators among fifth-grade students in three Bandung elementary schools. The overall mean score of 45.83, with the majority of students classified in the Poor category and none reaching Excellent, indicates that inferential reading competence in this sample was substantially below expected levels. The four indicators formed a clear difficulty gradient: anaphoric resolution was the strongest, followed by background knowledge activation and predictive reasoning at comparable levels, and retrospective synthesis as the weakest. Teacher and student data corroborated the test findings and further revealed that no systematic strategy for teaching inferential skills had been implemented in any of the three participating schools.

The study makes three contributions. First, it provides a dimension-specific empirical baseline of inferential reading ability at the elementary level, addressing a gap in the literature. Second, the four-indicator framework, integrating anaphoric resolution, background knowledge activation, predictive reasoning, and retrospective synthesis, offers educators a diagnostic vocabulary for identifying which particular inferential operations require targeted attention. Third, the triangulation of test, questionnaire, and interview data yields a multi-layered diagnostic profile that captures not only measured performance but also the perceptions of teachers and students.

Several factors may have contributed to the performance patterns observed, although the descriptive design of this study precludes causal attribution. The teacher interviews suggest that inferential reading skills had not been systematically taught in the participating schools, which represents a plausible instructional explanation. Other factors not measured in this study, including students' general language ability, family literacy environment, and variations in classroom reading practices, may also have played a role and warrant investigation in future research.

The implications of these findings are limited to the context of the participating schools and should not be generalized to Indonesian elementary education more broadly. Within that scope, the results suggest that reading instruction in settings comparable to the three schools studied may benefit from a more deliberate focus on inferential comprehension, particularly on retrospective synthesis and predictive reasoning. The cultivation of comprehension monitoring, which the interview data revealed to be largely absent among participating students, represents another potential area for instructional development. Future research should pursue validation of the inferential reading test with larger and more diverse samples, replication of the diagnostic approach in different geographic and socioeconomic

contexts, and intervention studies designed to test whether the difficulty gradient observed across the four indicators can be altered through targeted instruction.

REFERENCE

Ardiansyah, M., & Hardhienata, S. (2025). Modeling and Optimization of Teacher Innovativeness in Mathematics Instruction: A Study on Permanent Teachers at Private Vocational School Foundations in West Jakarta Using the POP SDM Method. *Formatif: Jurnal Ilmiah Pendidikan MIPA*, *15*(2), 59–70.

Attaprechakul, D. (2013). Inference Strategies to Improve Reading Comprehension of Challenging Texts. *English Language Teaching*, *6*(3), 82–91.

Badecker, W., & Straub, K. (2002). The processing role of structural constraints on interpretation of pronouns and anaphors. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(4), 748.

Cain, K., & Oakhill, J. (2007). *Children's comprehension problems in oral and written language: A cognitive perspective*. Guilford Press.

Chall, J. S. (1983). *Stages of reading development*. McGraw-Hill.

Corkery, F. (2014). *An analysis of year 8 poor comprehenders' responses to the PAT Reading Comprehension test: A thesis submitted in partial fulfilment of the requirements for the degree of Master of Literacy Education at Massey University, Palmerston North, New Zealand* [PhD Thesis, Massey University]. https://mro.massey.ac.nz/items/2b309e81-bf71-4287-adf7-f3980a89af29

Creswell, J. W. (2002). *Educational research: Planning, conducting, and evaluating quantitative* (Vol. 7). Prentice Hall Upper Saddle River, NJ.

Creswell, J. W., & Creswell, J. D. (2022). *Research design: Qualitative, quantitative, and mixed methods approach* (6th ed., Ed.). SAGE Publications.

Eilers, S., Tiffin-Richards, S. P., & Schroeder, S. (2019). The repeated name penalty effect in children's natural reading: Evidence from eye tracking. *Quarterly Journal of Experimental Psychology*, *72*(3), 403–412. https://doi.org/10.1177/1747021818757712

Elbro, C., & Buch-Iversen, I. (2013). Activation of Background Knowledge for Inference Making: Effects on Reading Comprehension. *Scientific Studies of Reading*, *17*(6), 435–452. https://doi.org/10.1080/10888438.2013.774005

Fox, E., & Alexander, P. A. (2014). Text comprehension: A retrospective, perspective, and prospective. In *Handbook of research on reading comprehension* (pp. 251–263). Routledge. https://api.taylorfrancis.com/content/chapters/edit/download?identifierName=doi&identifierValue=10.4324/9781315759609-22&type=chapterpdf

Hughson, A. L. (2003). *Cork and talk: The cognitive and perceptual bases of wine expertise*. University of Sydney (Australia). https://search.proquest.com/openview/2e4f1ee54392ff02c9b0a7ce6bcfb955/1?pq-origsite=gscholar&cbl=18750&diss=y

Hulme, C., & Snowling, M. J. (2011). Children's reading comprehension difficulties: Nature, causes, and treatments. *Current Directions in Psychological Science*, *20*(3), 139–142. https://doi.org/10.1177/0963721411408673

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge University Press.

Lata, K., Singh, P., & Dutta, K. (2021). A comprehensive review on feature set used for anaphora resolution. *Artificial Intelligence Review*, *54*(4), 2917.

Nation, K., Cocksey, J., Taylor, J. S. H., & Bishop, D. V. M. (2010). A longitudinal investigation of early reading and language skills in children with poor reading comprehension. *Journal of Child Psychology and Psychiatry*, *51*(9), 1031–1039. https://doi.org/10.1111/j.1469-7610.2010.02254.x

Oakhill, J., & Cain, K. (2019). Supporting reading comprehension development. *Perspectives on Language and Literacy*, *45*, 46–53.

Palumbo, A., & Sanacore, J. (2009). Helping Struggling Middle School Literacy Learners Achieve Success. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, *82*(6), 275–280. https://doi.org/10.3200/TCHS.82.6.275-280

Patterson, C. (2013). *The role of structural and discourse-level cues during pronoun resolution* [PhD Thesis, Universität Potsdam]. https://publishup.uni-potsdam.de/frontdoor/index/index/docId/6880

Rahmawati, L. E., Octaviani, P., Kusmanto, H., Nasucha, Y., & Huda, M. (2021). The accuracy of complex-procedures texts material in Bahasa Indonesia textbook for the first grade of senior high school. *Asian Journal of University Education (AJUE)*, *17*(1), 91–99.

Sanford, A. J., & Garrod, S. C. (1989). What, when, and how?: Questions of immediacy in anaphoric reference resolution. *Language and Cognitive Processes*, *4*(3–4), SI235–SI262. https://doi.org/10.1080/01690968908406369

Snow, C. E. (2002). *Reading for understanding: Toward an R&D program in reading comprehension*. RAND Corporation.

Trabasso, T., & Magliano, J. P. (1996). Conscious understanding during comprehension. *Discourse Processes*, *21*(3), 255–287. https://doi.org/10.1080/01638539609544959

van den Broek, P., Beker, K., & Oudega, M. (2015). Coherence inferences. In J. Oakhill, K. Cain, & C. Elbro (Eds), *Understanding and teaching reading comprehension: A handbook* (pp. 33–52). Routledge.

Wahyudi, R., & Rozimela, Y. (2024). The reciprocal teaching and reading interest impact on students' reading comprehension. *Elsya: Journal of English Language Studies*, *6*(3), 209–230.

Widiastuti, I. (2025). Assessing the Impact of Education Policies in Indonesia: Challenges, Achievement, and Future Direction. *Al-Ishlah: Jurnal Pendidikan*, *17*(2), 1955–1964.

Zhou, K., Chen, H., Ismail, M., & Yang, T. (2025). Assessment of Sustainable Development Levels, Country Disparities, and Dynamic Evolution Trends of ASEAN -10 Countries. *Sustainable Development*, *33*(6), 8547–8568. https://doi.org/10.1002/sd.70107

ACKNOWLEDGEMENT

The authors wish to thank the principals, teachers, and fifth-grade students of the three participating elementary schools in Zona D, Kota Bandung, for their generous cooperation and willingness to take part in this study. Gratitude is also extended to the expert reviewers who provided valuable feedback during the instrument validation process.

***Indra Suhendra (Corresponding Author)**
Department of Elementary Education, Faculty of Education,
Universitas Pendidikan Indonesia,
Jl. Dr. Setiabudi No.229, Isola, Kec. Sukasari, Kota Bandung, Jawa Barat 40154
Email: indrasuhendra@upi.edu

Wahyu Sopandi
Department of Elementary Education, Faculty of Education,
Universitas Pendidikan Indonesia,
Jl. Dr. Setiabudi No.229, Isola, Kec. Sukasari, Kota Bandung, Jawa Barat 40154
Email: wsopandi@upi.edu

Isah Cahyani
Department of Language Education, Faculty of Education,
Universitas Pendidikan Indonesia,
Jl. Dr. Setiabudi No.229, Isola, Kec. Sukasari, Kota Bandung, Jawa Barat 40154
Email: isahcahyani@upi.edu

Cepi Riyana
Department of Elementary Education, Faculty of Education,
Universitas Pendidikan Indonesia,
Jl. Dr. Setiabudi No.229, Isola, Kec. Sukasari, Kota Bandung, Jawa Barat 40154
Email: cheppy@upi.edu