

A Genre-Based Diagnostic Classroom Assessment of EFL Students' Reading Comprehension

Dedi Irwan

Universitas PGRI Pontianak, Indonesia

*Corresponding Author

Email: dediirwanphd@gmail.com

Abstract

Reading assessments in EFL classrooms frequently produce a single aggregate score that obscures the underlying sources of students' comprehension difficulties. This study advances a genre-based diagnostic classroom assessment designed to reveal how reading performance varies across text types and to identify the specific cognitive and discourse-level processes requiring instructional attention. Grounded in Systemic Functional Linguistics (SFL) and genre theory, the framework treats genre as the analytical unit, examining how staged organization and linguistic realization shape comprehension outcomes. A 20-item multiple-choice assessment was administered to 15 second-semester EFL students, with items systematically mapped to five reading sub-skills: identifying main ideas, locating supporting details, tracking references, making inferences, and interpreting vocabulary in context. Descriptive statistics were integrated with genre-sensitive diagnostic interpretation. Results indicate systematic variation across genres: students performed more successfully on descriptive and narrative texts (70%) than on recount and procedure texts (60%). Sub-skill analysis further revealed a marked imbalance between literal comprehension (main ideas, details) and discourse-level integration (reference resolution, inferencing, contextual vocabulary), with gaps exceeding 25 percentage points. These findings demonstrate that reading difficulty is structurally patterned rather than uniform, intensifying when texts require chronological sequencing, logical progression, and cohesive tracking. By disaggregating performance according to genre and sub-skill, the proposed framework exposes diagnostic patterns that remain concealed in conventional aggregate scoring. The study reframes genre not only as a pedagogical construct but as a diagnostic variable capable of structuring classroom reading assessment. Pedagogically, the results underscore the need for explicit instruction in genre staging, cohesive devices, and inferential reasoning, particularly in structurally demanding genres such as recount and procedure.

Keywords: *diagnostic assessment, EFL reading, genre-based assessment, reading comprehension*

INTRODUCTION

Reading comprehension in EFL contexts involves far more than vocabulary recognition and sentence decoding. Contemporary research emphasizes that comprehension requires the integration of linguistic knowledge, structural awareness, inferential reasoning, and discourse-level processing (Grabe, 2009). Learners must not only retrieve explicit information but also construct meaning by tracking references, making inferences, and recognizing how ideas are organized across a text (Duke & Pearson, 2002). Studies on word-to-text integration further show that syntactic processing and semantic association contribute significantly to both literal and inferential comprehension for L2 readers, even when vocabulary knowledge and working memory are controlled statistically (Alshehri & Zhang, 2022). A recent meta-analytic structural equation modeling study also demonstrates

that language comprehension explains more variance in reading outcomes than decoding skills, highlighting the importance of higher-order discourse and metalinguistic processes in L2 reading (Lee & Lee, 2024).

Despite this understanding, reading assessment in many EFL classrooms remains dominated by multiple-choice tests that produce a single overall score. Such assessments rarely explain why students struggle or which aspects of comprehension require instructional attention. From a formative assessment perspective, assessment should diagnose specific learning needs rather than merely measure performance (Black & Wiliam, 1998; Schildkamp et al., 2020; Jönsson, 2020). However, reading assessments seldom provide this diagnostic insight.

One important factor often overlooked in reading assessment is text structure. Research consistently shows that readers' awareness of how texts are organized significantly affects comprehension. Readers who recognize rhetorical and structural patterns comprehend more effectively because they can anticipate how information unfolds (Meyer & Ray, 2011). Meta-analytic evidence confirms that explicit instruction in text structure improves reading comprehension outcomes, especially in informational and expository texts (Bogaerds-Hazenberg, Evers-Vermeul, & van den Bergh, 2021; Hebert, Bohaty, Nelson, & Brown, 2016). These findings suggest that reading difficulty may arise not only from linguistic complexity but also from learners' unfamiliarity with discourse organization.

Genre theory, particularly within Systemic Functional Linguistics (SFL), provides a framework for understanding how texts organize meaning through staged, goal-oriented structures (Halliday & Matthiessen, 2014; Martin & Rose, 2008). Each genre, such as descriptive, narrative, recount, and procedure, possesses distinctive schematic structures and linguistic features that shape how meaning unfolds (Christie & Derewianka, 2008). While genre-based pedagogy has been widely applied in writing instruction (Nagao, 2019), and genre awareness has been shown to enhance reading comprehension (Arevalo, 2019; Gunawan, 2022; Putra, 2023), these insights have rarely informed the design of reading assessments.

At the same time, reading difficulties are frequently linked to specific sub-skills such as inferencing, reference tracking, and vocabulary interpretation rather than vocabulary knowledge alone (Afflerbach et al., 2011; Oakhill & Cain, 2012; Perfetti & Stafura, 2014). From an SFL perspective, reference is a key cohesive resource connecting meaning across clauses (Halliday & Matthiessen, 2014), and failure to resolve referential links often leads to comprehension breakdowns (McNamara et al., 1996). These challenges may become more pronounced in genres that require readers to track chronological sequences or logical steps.

Cognitive Load Theory provides an additional explanation for why some genres may be more difficult to comprehend. Tasks that require simultaneous processing of multiple elements can exceed working memory capacity (Paas, Renkl, & Sweller, 2020). Recount texts demand chronological tracking, while procedure texts require monitoring logical sequences of actions. Empirical evidence shows that reading tasks can impose varying cognitive load depending on structure and presentation (Brüggemann et al., 2023), suggesting that genre structure itself may contribute to comprehension difficulty.

Recent research in formative and diagnostic assessment further emphasizes that assessments should uncover specific strengths and weaknesses in reading processes rather than report aggregate scores (Bellido-García et al., 2026; Azmoode, 2025; Nuraini, Sodik, & Prastikawati, 2025). However, little research has explored how genre-based frameworks can be used to design such diagnostic reading assessments. Taken together, the literature indicates that reading comprehension is closely tied to text structure awareness, genre organization, inferencing, reference tracking, and cognitive processing demands. Yet these perspectives have rarely been integrated into a unified reading assessment framework

capable of revealing how students comprehend different genres. This study addresses that gap by proposing a genre-based diagnostic classroom assessment designed to make visible patterns of reading comprehension that remain hidden in conventional reading tests.

Specifically, this study seeks to answer the following questions:

- 1) How does students' reading comprehension vary across different text genres?
- 2) Which reading sub-skills (main idea, detail, reference, inference, vocabulary) present the greatest challenges?
- 3) How can a genre-based reading assessment function as a diagnostic tool for classroom instruction?

By integrating insights from reading theory, genre theory, cognitive load theory, and assessment research, this study offers a practical classroom-based framework for diagnosing reading comprehension in EFL contexts."

METHOD

Design

This study employed a genre-based diagnostic classroom assessment design to examine how students' reading comprehension varies across different text types. Rather than conceptualizing reading as a single measurable skill, the design assumes that comprehension is shaped by the social purpose, schematic structure, and linguistic realization of each genre, as articulated within Systemic Functional Linguistics (SFL) (Halliday & Matthiessen, 2014; Martin & Rose, 2008).

Within this perspective, texts are understood as staged, goal-oriented social processes. Consequently, readers are required to process meaning differently depending on how texts are organized and how communicative purposes are realized through staging and lexicogrammatical features. For this reason, multiple genres – descriptive, narrative, recount, and procedure – were deliberately selected as the unit of assessment to capture patterns of comprehension that would remain concealed in a single aggregate reading score. Positioned as a genre-based diagnostic classroom assessment, the study does not merely rank student performance. Instead, it interprets performance variations in relation to genre structure and linguistic demands, enabling the identification of specific genres and reading processes that require instructional attention. This orientation aligns with classroom assessment perspectives that prioritize diagnostic and pedagogical value over statistical generalization (Black & Wiliam, 1998; Jönsson, 2020).

Participants and Context

The participants were 15 second-semester students enrolled in an English Education program at a private university in Indonesia. A purposive classroom-based selection approach was adopted, as the study was designed as a diagnostic investigation within an intact instructional setting rather than for statistical generalization. The selected class had completed foundational courses in reading and grammar and had prior exposure to genre-based texts commonly included in the EFL curriculum (descriptive, narrative, recount, and procedure). This ensured that students possessed sufficient background familiarity with the genres under investigation, thereby allowing the assessment to focus on diagnostic comprehension patterns rather than first-time genre exposure effects. The use of an intact class aligns with principles of classroom assessment research, which prioritize ecological validity and instructional relevance over random sampling (Bennett, 2011; Jönsson, 2020).

Instrument: Genre-Based Reading Assessment

The instrument consisted of a 20-item multiple-choice reading comprehension test constructed from four different genres:

Table 1. Instrument Blueprint

Genre	Items	Communicative Purpose
Descriptive	5	Describing objects/phenomena
Narrative	5	Telling imaginative events
Recount	5	Retelling past events chronologically
Procedure	5	Giving instructions and steps

These genres were selected because they represent distinct schematic structures and linguistic features commonly found in EFL curricula (Martin & Rose, 2008). Each text reflected authentic genre characteristics, enabling diagnostic interpretation linked to genre staging and lexicogrammatical patterns.

Mapping Questions to Reading Skills

To enhance the diagnostic function, items were mapped to key reading sub-skills widely discussed in reading research (Grabe, 2009):

Table 2. Map of Reading Skill Questions

Reading Skill	Purpose
Identifying main ideas	Global comprehension
Locating supporting details	Local comprehension
Recognizing references	Cohesion tracking
Making inferences	Inferential comprehension
Understanding vocabulary in context	Lexical inference

This mapping allowed the study to identify not only genre-related differences but also cognitive processes that posed difficulty for students.

Genre Analysis Framework

Each text was analyzed using an SFL genre framework focusing on:

1. Generic structure (e.g., Orientation–Complication–Resolution; Goal–Materials–Steps)
2. Linguistic features (tense, connectors, reference items, imperative verbs)
3. Social purpose of the genre

This analysis provided the basis for interpreting why certain genres were more challenging for students (Christie & Derewianka, 2008; Halliday & Matthiessen, 2014).

Data analysis

Data analysis was conducted through a systematic two-stage procedure designed to ensure both statistical clarity and theoretical depth. In the first stage, descriptive statistical analysis was performed to generate the overall mean score, the mean score for each genre, and the distribution of performance across specific reading skills. This stage provided a comprehensive overview of students' reading achievement patterns and allowed for the identification of variation across genres and skill domains.

In the second stage, a genre-based diagnostic interpretation was undertaken to examine performance patterns in relation to the structural characteristics of each genre. Students' difficulties were analyzed by considering the linguistic complexity and cognitive demands embedded within different genre types. The interpretation was further grounded in established genre theory and contemporary reading comprehension research, enabling a theoretically informed explanation of observed disparities. By integrating descriptive statistics with diagnostic interpretation, this two-stage approach moved beyond mere score reporting and offered a structured explanation for why comprehension levels varied across genres (Paas et al., 2020).

Ethical Considerations

Ethical procedures were carefully observed in conducting this classroom-based study. Prior to data collection, students were informed verbally and in writing that the reading assessment would be used for research purposes in addition to instructional diagnosis. Participation was voluntary, and students were assured that their decision to participate or withdraw would not affect their course grades. Written informed consent was obtained from all participants. No personal identifiers were recorded; student responses were anonymized using numerical coding to ensure confidentiality. The assessment was administered as part of normal classroom activities, minimizing disruption and avoiding any additional psychological burden. The study adhered to institutional research ethics principles concerning voluntary participation, confidentiality, and responsible data use. All collected data were used solely for academic research and pedagogical development purposes.

RESULTS

Variation of Reading Comprehension Across Genres

The first analytical focus of this study concerns how students' reading comprehension varied across different text genres. Analysis of the genre-based assessment demonstrates that comprehension performance was not evenly distributed across the four text types. Although the overall mean score across all genres was 65%, genre-level disaggregation reveals systematic variation. Students achieved higher mean scores in descriptive and narrative texts (70% each), while performance in recount and procedure texts was lower (60% each). The ten-point gap between these two clusters indicates that comprehension outcomes were sensitive to genre type rather than uniformly distributed. This difference was consistent across participants, suggesting that the pattern reflects structural characteristics of the texts rather than isolated individual variation.

Descriptive and narrative genres formed a higher-performing cluster, whereas recount and procedure texts constituted a lower-performing cluster. Students who demonstrated relatively strong performance in descriptive and narrative texts did not show proportional improvement in recount and procedure texts. This consistency across participants suggests that the observed differences were systematic and related to genre-level demands rather than incidental fluctuations in individual performance.

The distribution pattern indicates that certain text types were cognitively more manageable, while others introduced additional processing demands. The ten-percentage-point gap, while moderate at the aggregate level, becomes more meaningful when considered in relation to the overall mean of 65%. It suggests that comprehension was genre-dependent and that performance shifted in predictable ways depending on text organization and communicative purpose.

Patterns of Performance Across Reading Sub-Skills

The second analytical focus addresses which reading sub-skills presented the greatest challenges. Item-level analysis across five sub-skills—identifying main ideas, locating supporting details, recognizing references, making inferences, and interpreting vocabulary in context—reveals a clear differentiation between literal and discourse-level comprehension. Students demonstrated relatively strong performance in tasks requiring retrieval of explicitly stated information. The mean score for identifying main ideas was 2.93 out of 4 items (73%), and for locating supporting details was 3.80 out of 5 items (76%). These results indicate that students were generally capable of extracting directly presented information and locating key points within texts.

In contrast, performance declined substantially in tasks requiring integration beyond sentence-level meaning. Reference recognition yielded a mean score of 1.47 out of 3 items (49%), inferential questions produced an average of 0.93 out of 2 items (47%), and

vocabulary in context showed the lowest performance, with an average of 0.40 out of 1 item (40%). The gap between literal comprehension (above 70%) and discourse-level processing (below 50%) exceeds 25 percentage points. This discrepancy signals a pronounced imbalance between students' ability to retrieve explicit information and their ability to integrate, infer, and resolve meaning across textual units. The pattern was consistent across participants, indicating that the observed weaknesses were not isolated to a small subgroup but reflected a broader processing challenge.

The results therefore identify reference tracking, inferencing, and contextual vocabulary interpretation as the most significant areas of difficulty. These sub-skills required students to connect information across clauses, maintain continuity of participants, and construct meaning beyond explicitly stated content, which corresponded with lower levels of accuracy.

Interaction Between Genre and Skill Demands

Closer examination of genre-level and skill-level data together reveals a structural interaction pattern. The lower-performing genres—recount and procedure—contained a greater concentration of items that required sequential tracking, reference resolution, and inferential integration. In contrast, descriptive and narrative texts included a higher proportion of items centered on explicit information retrieval. This alignment suggests that comprehension difficulty intensified when texts demanded sustained monitoring of relationships across clauses, particularly in sequences of events or procedural steps. When meaning construction required continuous updating of participants, actions, or logical progression, performance declined noticeably.

Further inspection indicates that this interaction was not driven by isolated items but emerged as a consistent pattern across the dataset. Students who maintained strong performance in literal tasks within descriptive and narrative genres did not maintain equivalent performance when confronted with structurally demanding recount and procedure texts. This consistency suggests that comprehension breakdown occurred when genre structure activated discourse-level processing demands in which students showed lower proficiency. Taken together, the findings indicate that reading performance was shaped by the interaction between genre organization and the cognitive processes required to navigate that organization. Rather than reflecting isolated weaknesses, the results point to a patterned relationship between text structure and skill activation.

Diagnostic Value of the Genre-Based Assessment

The third analytical focus concerns the extent to which the genre-based reading assessment functioned diagnostically within classroom practice. When interpretation is limited to the overall mean score of 65%, students' reading proficiency appears moderately developed, offering only a general indication of performance. However, disaggregation by genre and sub-skill reveals a far more differentiated comprehension profile. Three diagnostic signals emerge clearly. First, comprehension varies systematically across genres, with recount and procedure texts presenting greater difficulty than descriptive and narrative texts. Second, students exhibit relative strength in literal comprehension but substantial weakness in discourse-level integration. Third, specific vulnerabilities are concentrated in reference tracking, inferential reasoning, and contextual vocabulary interpretation. The assessment therefore generates a structured map of strengths and weaknesses rather than a single aggregate indicator. By revealing where comprehension breakdown occurs—both in terms of genre and cognitive process—the framework enables more precise identification of instructional priorities. The differentiated pattern of results demonstrates that reading performance is shaped by both text type and the cognitive demands embedded within each genre.

DISCUSSION

With regard to the first research question concerning how reading comprehension varies across genres, the findings demonstrate that comprehension outcomes are clearly genre-sensitive rather than uniform across text types. The consistent ten-point performance gap between descriptive/narrative texts (70%) and recount/procedure texts (60%) indicates that structural organization exerts a measurable influence on comprehension. This difference cannot be dismissed as incidental fluctuation; instead, it reflects patterned variation aligned with genre-specific demands.

Descriptive texts, which typically revolve around identification and attribute elaboration, offer relatively stable topic continuity. Meaning development in such texts tends to unfold through additive description rather than through structural progression. Similarly, narrative texts provide recognizable staging through orientation, complication, and resolution, allowing readers to anticipate how events will develop. These predictable patterns appear to reduce structural ambiguity and enable readers to allocate cognitive resources toward consolidating meaning rather than constantly reorganizing mental representations. In contrast, recount and procedure texts impose additional structural constraints. Recount texts require readers to maintain chronological sequencing while tracking participant continuity across temporal shifts. Procedure texts demand logical monitoring of steps, goals, and conditional progression. These structures require readers to update representations continuously rather than accumulate information linearly. From a cognitive load perspective (Paas et al., 2020), such concurrent demands may increase processing burden when readers must simultaneously decode lexical content, maintain cohesion, and monitor structural progression. The observed performance gap therefore suggests that comprehension difficulty intensifies when structural monitoring becomes central to meaning construction.

Concerning the second research question, which examined which reading sub-skills posed the greatest challenges, the results reveal a pronounced imbalance between literal retrieval and discourse-level integration. Students performed relatively well in identifying main ideas (73%) and supporting details (76%), indicating that explicit information retrieval was not a major obstacle. However, performance declined substantially in reference recognition (49%), inferencing (47%), and contextual vocabulary interpretation (40%). The magnitude of this gap—exceeding 25 percentage points—signals a systematic divergence between surface-level and integrative processing.

Within systemic functional linguistics, reference functions as a cohesive mechanism that sustains continuity of participants and events across clauses (Halliday & Matthiessen, 2014). When referential chains are not successfully resolved, coherence becomes fragmented. Empirical research further demonstrates that text cohesion interacts with background knowledge and influences comprehension depth (McNamara et al., 1996; Smith et al., 2021). Inferencing similarly requires readers to construct meaning beyond explicitly stated information, integrating dispersed textual cues into coherent representations (Oakhill & Cain, 2012; Shi, 2023). Vocabulary interpretation in context further intensifies processing demands when lexical meaning must be inferred rather than recognized directly (Nation, 2013; Perfetti & Stafura, 2014). The convergence of weaknesses across these areas indicates that students' comprehension remains predominantly literal, while discourse-level meaning construction remains underdeveloped.

Importantly, the interaction between genre and sub-skill demands provides a more nuanced explanation than either dimension alone. Recount and procedure genres activate precisely those processes—sequential tracking, cohesive resolution, inferential bridging—in which students demonstrated lower accuracy. The alignment between structural complexity and skill vulnerability suggests that comprehension difficulty arises not from isolated lexical deficiency but from the interaction between genre staging and cognitive

processing capacity. In other words, genre functions as a trigger that activates specific cognitive demands, and performance declines when those demands exceed learners' integration abilities.

In response to the third research question concerning the diagnostic function of a genre-based assessment, the findings illustrate how disaggregating performance by genre and sub-skill produces substantially richer instructional insight than relying on an aggregate score of 65%. An overall mean suggests moderate proficiency but conceals structural variability. By contrast, genre-sensitive analysis reveals patterned strengths and vulnerabilities that can inform instructional decision-making. Conventional reading assessments often emphasize aggregate accuracy without situating performance within discourse organization. In many cases, sub-skill scores are reported independently of text type, which limits interpretive depth. By treating genre as the analytical unit, the present framework captures how comprehension fluctuates according to communicative purpose and staged organization (Martin & Rose, 2008; Christie & Derewianka, 2008). This alignment strengthens the diagnostic validity of the assessment because it connects observed performance to identifiable structural demands. Furthermore, the approach resonates with formative assessment principles that emphasize using evidence to guide instructional adaptation rather than merely classify performance (Black & Wiliam, 1998; Jönsson, 2020; Schildkamp et al., 2020).

The pedagogical implications emerging from these findings extend beyond isolated skill practice. First, explicit genre orientation becomes a strategic necessity rather than a supplementary activity. When students understand the staging and purpose of recount and procedure texts before reading, they may allocate attention more efficiently and reduce unnecessary cognitive load. Second, reference tracking should be integrated as a recurring classroom routine. Structured activities that require students to trace cohesive chains across sentences can strengthen discourse-level continuity and mitigate fragmentation. Third, inferencing practice should move from sporadic questioning to systematic scaffolding, encouraging students to articulate causal, temporal, and logical relationships embedded in texts.

Additionally, vocabulary instruction should shift toward contextual integration rather than isolated memorization. In structurally demanding genres, key lexical items often signal sequence, condition, or action. Helping students recognize these functional cues can enhance procedural and chronological comprehension. Importantly, these instructional responses are directly grounded in the diagnostic profile generated by the assessment. Rather than applying generalized comprehension exercises, teachers can target specific genre-skill intersections identified through evidence. Beyond immediate classroom application, the study contributes to a broader reconceptualization of diagnostic reading assessment. Genre is positioned not merely as a pedagogical category but as a diagnostic variable capable of structuring performance evidence. When assessment is aligned with discourse structure, it reveals how text organization shapes cognitive activation patterns. This alignment bridges genre theory, cognitive processing perspectives, and formative assessment practice within a unified framework.

Overall, the study demonstrates that reading comprehension in EFL contexts emerges from the interaction between genre structure, cohesive processing, inferential reasoning, and cognitive resource management. By situating diagnostic assessment within this interactional model, the genre-based framework advances a structurally grounded approach to classroom reading evaluation. It moves beyond aggregate scoring toward a nuanced profiling system that informs targeted instructional design. In doing so, it strengthens the conceptual and practical integration of genre theory and diagnostic assessment within EFL reading research.

Limitations and Future Research

Several limitations should be acknowledged. First, the study involved a small sample of 15 students from a single classroom. While appropriate for a diagnostic classroom-based investigation, this limits statistical generalization beyond the immediate context. The findings should therefore be interpreted as context-specific instructional insights rather than population-level conclusions. Second, each genre and reading sub-skill was represented by a limited number of items. Although the instrument was designed for practical classroom use, a larger item pool could enhance reliability and provide more precise diagnostic differentiation, particularly for inferencing and reference tracking. Third, comprehension was measured exclusively through multiple-choice responses.

While this format enables structured comparison across genres, it does not capture real-time cognitive processing during reading. Moreover, assessment mode and presentation format may influence experienced cognitive load and performance (Brüggemann et al., 2023). Future studies could incorporate process-oriented methods or compare assessment modes to examine how genre structure interacts with cognitive load under different testing conditions. Finally, the study did not examine the instructional impact of the diagnostic findings. Future research should replicate the framework with larger samples and test whether targeted genre-focused interventions lead to measurable improvements in discourse-level comprehension.

CONCLUSION

This study examined how reading comprehension varies across genres and how a genre-based classroom assessment can function diagnostically in EFL contexts. The findings demonstrate that comprehension is not uniform across text types. Descriptive and narrative texts were more accessible to students than recount and procedure texts, indicating that genre structure significantly shapes comprehension outcomes. Furthermore, skill-level analysis revealed a clear imbalance between literal comprehension and discourse-level integration, with students performing relatively well in identifying main ideas and supporting details but struggling in reference tracking, inferencing, and contextual vocabulary interpretation. These findings confirm that reading difficulty in EFL classrooms is not merely a matter of vocabulary deficiency but is closely linked to structural processing demands embedded within genre staging. Texts requiring sustained chronological or procedural tracking intensify the need for referential resolution and inferential reasoning, exposing specific areas of vulnerability. By disaggregating performance according to genre and sub-skill, the study demonstrates that a genre-based assessment framework provides a more precise diagnostic profile than aggregate scoring alone. Theoretically, this study contributes to reading assessment research by integrating genre theory, cognitive processing considerations, and formative diagnostic principles within a single classroom-based framework. Methodologically, it positions genre not only as a pedagogical construct but as an analytical unit for assessment. Practically, the findings suggest that EFL reading instruction should incorporate explicit attention to genre staging, cohesive devices, and inferential reasoning routines, particularly in structurally demanding genres such as recount and procedure. Overall, the study advances a discourse-sensitive diagnostic model that bridges assessment practice and instructional design. By foregrounding genre structure as a central variable in comprehension profiling, it offers a theoretically grounded and pedagogically actionable approach to strengthening reading development in EFL classrooms.

REFERENCES

- Afflerbach, P., Pearson, P. D., & Paris, S. G. (2011). Clarifying differences between reading skills and reading strategies. *The Reading Teacher*, 61(5), 364–373.

- <https://doi.org/10.1598/RT.61.5.1>
- Alshehri, M., & Zhang, D. (2022). Word-to-text integration components in second language (L2) reading comprehension. *Frontiers in Education*, 7, 926663. <https://doi.org/10.3389/educ.2022.926663>
- Bellido-García, R., Lozano-Gonzalez, B., Cruzata-Martínez, A., Rodríguez-Huamán, M., & Rejas-Borjas, G. (2026). Formative assessment and its contribution to the improvement of reading comprehension in a public school in Lima. *European Public & Social Innovation Review*, 11, 1–18. <https://doi.org/10.31637/epsir-2026-1956>
- Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice*, 18(1), 5–25. <https://doi.org/10.1080/0969594X.2010.513678>
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5(1), 7–74. <https://doi.org/10.1080/0969595980050102>
- 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>
- Brüggemann, T., Ludewig, U., Lorenz, R., & McElvany, N. (2023). Effects of mode and medium in reading comprehension tests on cognitive load. *Computers & Education*, 192, 104649. <https://doi.org/10.1016/j.compedu.2022.104649>
- Christie, F., & Derewianka, B. (2008). *School discourse: Learning to write across the years of schooling*. Continuum.
- Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. Cambridge University Press.
- Halliday, M. A. K., & Matthiessen, C. M. I. M. (2014). *Halliday's introduction to functional grammar* (4th ed.). Routledge. <https://doi.org/10.4324/9780203783771>
- Hebert, M., Bohaty, J. J., Nelson, J. R., & Brown, J. A. (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>
- Jin, G., Li, C., & Sun, Y. (2024). Developing a genre-based research article reading module for undergraduate students. *Journal of English for Academic Purposes*, 72, 101448. <https://doi.org/10.1016/j.jeap.2024.101448>
- Jönsson, A. (2020). Definitions of formative assessment need to make a distinction between a psychometric understanding of assessment and evaluative judgment. *Frontiers in Education*, 5, 2. <https://doi.org/10.3389/educ.2020.00002>
- Lee, H., & Lee, J. H. (2024). Extending the simple view of reading in second and foreign language learning: A meta-analytic structural equation modeling approach. *Review of Educational Research*, 94(4), 467–500. <https://doi.org/10.3102/00346543231186605>
- Lestari, T., & Damayanti, I. L. (2022). Reading to learn (R2L) pedagogy: Teaching reading comprehension to a young English language learner. *JEES (Journal of English Educators Society)*, 7(1), 1558–1569. <https://doi.org/10.17977/um064v2i112022p1558-1569>
- Martin, J. R., & Rose, D. (2008). *Genre relations: Mapping culture*. Equinox.
- McNamara, D. S., Kintsch, E., Songer, N. B., & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction*, 14(1), 1–43. https://doi.org/10.1207/s1532690xc1401_1
- Meyer, B. J. F., & Ray, M. N. (2011). Structure strategy interventions: Increasing reading comprehension of expository text. *Reading Research Quarterly*, 46(2), 127–151. <https://doi.org/10.1598/RRQ.46.2.3>
- Nation, I. S. P. (2013). *Learning vocabulary in another language* (2nd ed.). Cambridge University Press.

- Paas, F., Renkl, A., & Sweller, J. (2020). Cognitive load theory and instructional design: Recent developments. *Current Directions in Psychological Science*, 29(4), 394–402. <https://doi.org/10.1177/0963721420922183>
- Pickren, S. E., Stacy, M., Del Tufo, S. N., Spencer, M., & Cutting, L. E. (2022). The contribution of text characteristics to reading comprehension: Investigating the influence of text emotionality. *Reading Research Quarterly*, 57(2), 649–667. <https://doi.org/10.1002/rrq.431>
- Purwaningsih, L. (2024). A correlation study on cohesive device, vocabulary mastery and reading skill at eleventh grade students of SMK Maospati academic year 2023/2024. *Eduscotech*, 5(1). <https://journal.udn.ac.id/index.php/eduscotech/article/view/350>
- Schildkamp, K., van der Kleij, F. M., Heitink, M. C., Kippers, W. B., & Veldkamp, B. P. (2020). Formative assessment: A systematic review of critical teacher prerequisites for classroom practice. *International Journal of Educational Research*, 103, 101602. <https://doi.org/10.1016/j.ijer.2020.101602>
- Shi, Y. (2023). How inference making mediates the relation between vocabulary knowledge and reading comprehension for Chinese EFL middle school students. *English Teaching*, 78(3), 85–104. <https://doi.org/10.15858/engtea.78.3.202309.85>
- Smith, R., Snow, P., Serry, T., & Hammond, L. (2021). The role of background knowledge in reading comprehension: A critical review. *Reading & Writing Quarterly*, 37(3), 214–240. <https://doi.org/10.1080/02702711.2021.1888348>
- Xuan, Q., Cheung, A., & Sun, D. (2022). The effectiveness of formative assessment for enhancing reading achievement in K-12 classrooms: A meta-analysis. *Frontiers in Psychology*, 13, 990196. <https://doi.org/10.3389/fpsyg.2022.990196>