Vol. 7, No. 1, 2026, pp. 290-304
DOI: https://doi.org/10.29210/992260

Contents lists available at Elora Center

Lentera Negeri

Journal homepage: https://lentera.eloracenter.org/lentera

Exploring students' critical reading development through digital texts: a mixed methods study in indonesian efl context

Zulhidah Zulhidah*)
Universitas Islam Negeri Sultan Syarif Kasim Riau, Pekanbaru, Indonesia

Article Info

Article history:
Received Mar 12th, 2026
Revised Jun 4th, 2026
Accepted Jun 9th, 2026

Keyword:
Critical digital literacy
Critical reading strategies
Indonesian efl undergraduates
Source credibility evaluation

ABSTRACT

Critical reading has become a decisive academic competence in an information environment dominated by digital texts, yet many learners of English as a Foreign Language (EFL) read online material passively rather than critically. Because a single method can establish either how well students read or why they read as they do, but not both, this study adopted an explanatory sequential mixed methods design, pursuing three linked objectives: a quantitative objective to measure the level of students' critical reading across five dimensions, a qualitative objective to explain the strategies underlying that level, and an integrative objective to generate meta-inferences connecting the two. In Phase 1, 20 undergraduates in one English Education program at Universitas Islam Negeri Sultan Syarif Kasim Riau completed a 10-item Likert questionnaire whose content validity was confirmed by three expert reviewers (scale-level CVI = .92) and whose internal consistency was good (Cronbach's α = .86); responses were analyzed using descriptive statistics. In Phase 2, 10 students selected for maximum variation on their scores took part in semi-structured interviews analyzed thematically following Braun and Clarke, in which 38 initial codes were collapsed into six subthemes and two themes. Quantitatively, students rated themselves consistently high (composite M = 4.52, SD = 0.50). Qualitatively, two themes strategic critical engagement with digital texts and evaluative and reflective reading practices accounted for the scores. Integration through a joint display showed confirmation on four dimensions and, importantly, expansion on the lowest-rated dimension: the qualitative data revealed that students detected bias by comparing across texts rather than analyzing a single text, information the questionnaire alone could not yield. Because the design is cross-sectional and relies on self-report, the findings describe students' attained competence and reported strategies at one point in time rather than change over time or a causal effect of digital texts. The study contributes a worked example of strand integration in critical reading research and identifies within-text bias analysis as a clear instructional priority.

Corresponding Author:
Zulhidah Zulhidah,
zulhidah@uin-suska.ac.id

Introduction

The way university students read has changed profoundly over the past decade. Reading now happens predominantly on screens, and the texts students encounter are no longer curated by editors and publishers but flow continuously and unfiltered from websites, online journals, blogs, and social media platforms. In this environment the quality, intent, and reliability of what students read vary enormously, and the ability to

read critically, that is, to question, evaluate, and synthesize rather than simply absorb, has become one of the most consequential academic competencies a learner can possess. The urgency of this competence is heightened by what has been described as an era of technological and informational disruption, in which learners are continuously exposed to vast quantities of content of uneven trustworthiness and must decide for themselves what to believe (Khusniyah, 2024). For learners of English as a Foreign Language (EFL), the demand is compounded by a linguistic dimension, since they must perform these critical operations in a language that is not their first.

Before reviewing the evidence, it is useful to be precise about the construct at stake. Following Wallace (2003), critical reading is understood here not as a single skill but as a cluster of interrelated operations through which a reader interrogates a text rather than receiving it: identifying main ideas and arguments, analyzing how ideas relate and separating claims from evidence, evaluating the credibility of sources, recognizing bias and unstated assumptions, and drawing warranted conclusions and inferences. These operations have long been treated as cultivable rather than innate, and as foundational to building learners' knowledge base in language education (Mulumba, 2016); critical literacy pedagogies have correspondingly been shown to enhance EFL reading by teaching learners to question texts (Litim & Bouguebs, 2025).

Equally important is what is meant by development. This study uses the term in the sense of an attained, cultivated state of competence and a repertoire of strategies that learners have built through sustained, habitual engagement with digital English texts, rather than in the sense of measured change across time. Because the design is cross-sectional, it documents the level students have reached and the strategies they currently deploy; it does not, and cannot, track growth longitudinally or isolate the causal contribution of any single factor. Making this scope explicit at the outset guards against overclaiming and frames the digital text not as a cause but as an affordance whose value depends on how learners and teachers take it up.

That affordance has a theoretical basis. New literacies scholarship argues that reading online is not simply print reading transferred to a screen but a partly distinct practice, because digital texts are multimodal and non-linear and require readers to locate, evaluate, and synthesize across sources as part of comprehension itself (Leu et al., 2013). Empirical work on online reading comprehension similarly shows that source evaluation and cross-text synthesis are integral to, not separate from, understanding digital material (Coiro, 2011). These features expand what reading involves: a reader must distribute attention across competing elements, decide which links to follow, and reassemble meaning from fragments scattered across several sources. For EFL learners already managing the cognitive load of a second language, such demands can either overwhelm comprehension or, when handled strategically, become resources for deeper engagement. This is why the critical reading of digital texts cannot be assumed to transfer automatically from the skills learners build with linear print.

Set against this construct, the existing literature falls into three uneven strands, and reading them together reveals a tension rather than a settled picture. A first strand documents the broad importance of digital literacy for EFL reading: a systematic review confirms that these competences are central to contemporary language learning while noting that the field remains fragmented in how it conceptualizes and measures them (Luthfia, 2025), and scholars increasingly argue that twenty-first century literacy must extend beyond linguistic proficiency toward critical and digital empowerment (Ismael et al., 2026), a claim reinforced by the spread of brief, persuasive, multimodal online genres that demand scrutiny (Javahery et al., 2026; Afnan, 2023). A second strand reports that the specific ability most distinctive of digital reading, source evaluation, is frequently weak: studies of EFL learners have repeatedly found uneven capacity to judge online sources (Silvhiany et al., 2021), which has prompted purpose-built assessment instruments (Silvhiany & Huzaifah, 2021) and evidence that teachers' own practices shape whether such competence develops (Rini & Nabhan, 2023). A third strand, by contrast, reports comparatively strong or improvable critical engagement: multimodal scaffolding supports critical reading among Indonesian learners (Puspitasari et al., 2025), responsive pedagogy nurtures critical thinking in digital EFL reading (Iswati et al., 2025), students view online articles positively as critical-reading material (Ramadhini & Wahyuni, 2026), and structured intervention with digital books can raise critical reading (Hidayanto et al., 2026).

The inconsistency between the second and third strands is the crux. Some studies find learners poor at the very source-evaluation work that digital reading most requires, while others find critical engagement that is robust or readily strengthened, and research on Indonesian EFL students shows that research competence, critical thinking, and digital literacy do not always advance together (Indah et al., 2022). The medium itself does not resolve the matter, since comparative reviews show that strategies and comprehension differ between print and digital reading without favoring either uniformly (Al-Amrani, 2022; Rustambekova, n.d.), that the effect of medium interacts with learners' critical thinking dispositions (Rahayu et al., 2023), and that

digital environments scaffold reading only when integrated deliberately (Asia et al., 2024; Mistang et al., 2025). What the field lacks is not more measurement or more description in isolation, but studies that hold the two together long enough to see whether a measured level actually corresponds to documented practice.

This is precisely the gap a mixed methods design is built to close, and the rationale is substantive rather than merely procedural. Quantitative studies establish how well students perform but leave the reasoning opaque; qualitative studies illuminate practice without anchoring it to a measured level. When a sample rates itself uniformly high, a single quantitative strand cannot distinguish genuine competence from inflated self-assessment, and a single qualitative strand cannot say whether the strategies described are typical of high or low performers. Integration supplies what neither can alone: it tests whether self-rated competence is corroborated by concrete, behaviorally specific accounts, and it can reveal where the two diverge. In mixed methods terms, this is the value of complementarity and development across strands (Greene, 2007; Bryman, 2006; Teddlie & Tashakkori, 2009), and it is the logic that earlier work combining metacognitive measurement of online reading strategies with qualitative inquiry has exploited (Marboot et al., 2020).

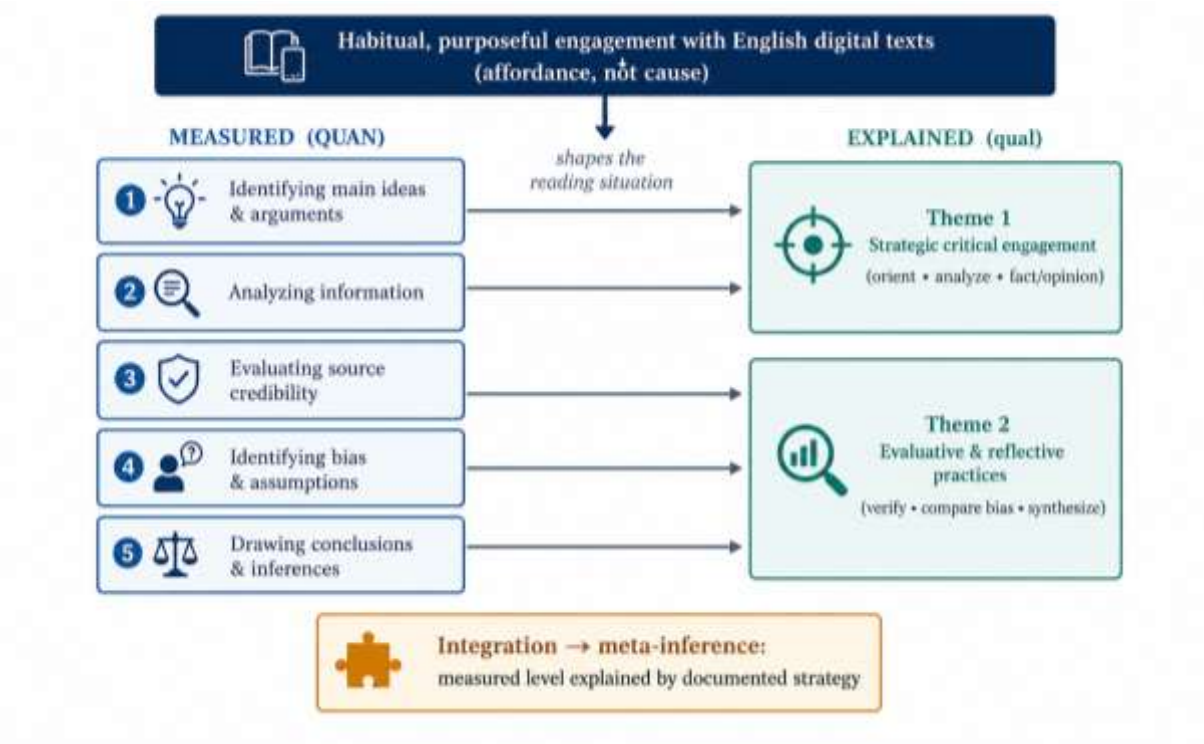

Figure 1. Digital Engagement Shapes Critical Reading

Figure 1 sets out the conceptual framework that organizes the study. Habitual engagement with English digital texts is positioned as an affordance that shapes a reading situation; within that situation, five measurable dimensions of critical reading are linked, through integration, to the reading strategies that the qualitative phase documents. The framework makes the analytic claim of the study explicit: the level recorded on the left should be explicable by the strategies recovered on the right, and the point of integration is to test and interpret that correspondence rather than to assume it.

Accordingly, the study pursues three logically connected objectives. The quantitative objective is to measure the level of students' critical reading across the five dimensions. The qualitative objective is to explain, in students' own accounts, the strategies that underlie that level. The integrative objective, which depends on the first two, is to generate meta-inferences that link measured competence to documented practice and to locate any dimension where the two do not simply agree. These objectives are expressed in three research questions. First, what is the level of students' critical reading through digital texts across the five dimensions? Second, what strategies do students employ when reading English digital texts critically? Third, how do students' accounts of those strategies explain, and where do they complicate, the levels identified quantitatively?

Method

Research Design

This study employed an explanatory sequential mixed methods design, a two-phase approach in which quantitative data are collected and analyzed first and qualitative data are then collected to explain and elaborate the quantitative results (Creswell & Plano Clark, 2018; Teddlie & Tashakkori, 2009). The design was chosen because the research questions required both a measurement of the level of students' critical reading and an explanation of the reasoning behind that level, and because integration of the two strands was expected to yield understanding unavailable from either alone (Fetters et al., 2013). Following standard notation, the design is represented as QUAN → qual, with the quantitative strand dominant and prior in time and the qualitative strand supplementary and explanatory. The two strands were connected at the point of participant selection, where quantitative results guided the choice of interviewees, and were integrated at the interpretation stage through a joint display (Guetterman et al., 2015). Figure 2 summarizes the procedure.

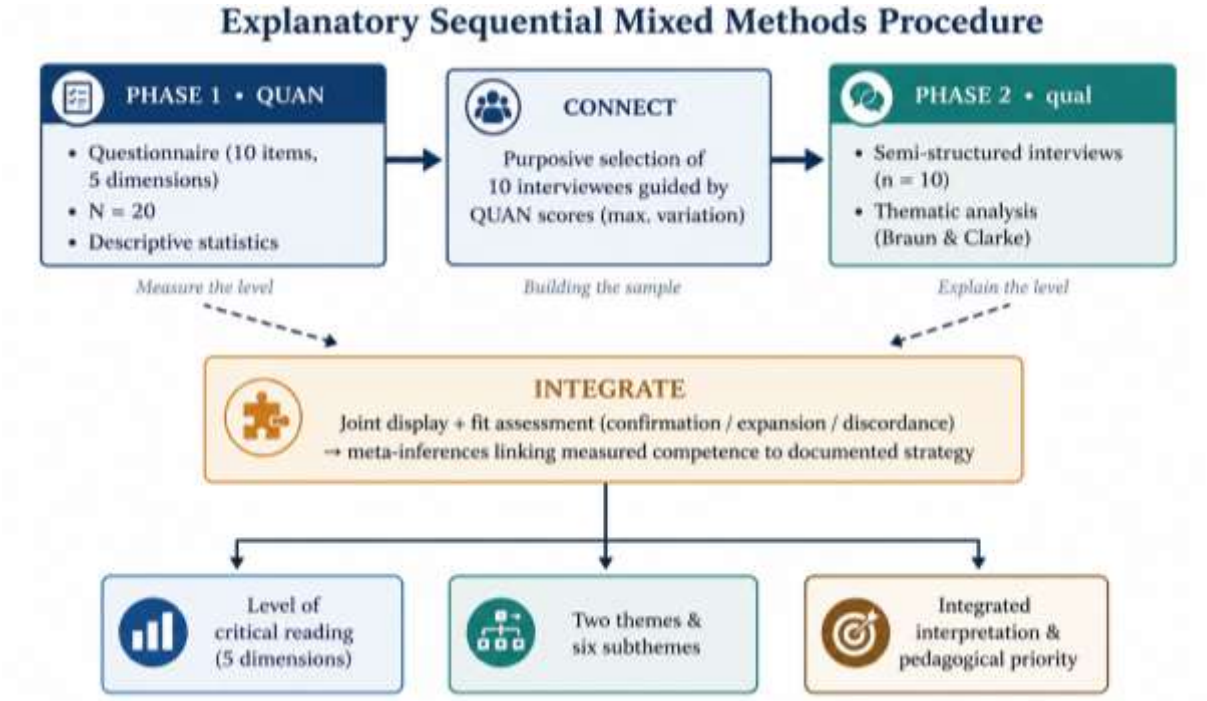

Figure 2. Explanatory Sequential Mixed Methods Procedure

Participants and Sampling

Participants were undergraduate students of the English Education Department in the Faculty of Education and Teacher Training at Universitas Islam Negeri Sultan Syarif Kasim Riau, Pekanbaru, all of whom read English digital texts, including online articles, e-books, and academic web sources, as a routine part of their coursework. In the quantitative phase, 20 students completed the questionnaire. This sample size was not intended to support inferential generalization to the wider EFL population; consistent with the exploratory, descriptive function of the first strand in an explanatory sequential design, it was sufficient to establish the level and dispersion of scores within this program and to provide the score distribution from which information-rich interviewees could be drawn. The limits this places on the quantitative claims are acknowledged directly in the discussion.

For the qualitative phase, 10 of the 20 students were chosen through purposive sampling using a maximum-variation logic. Three explicit criteria governed selection: spread across the observed range of total scores, so that relatively lower as well as higher scorers were represented; deliberate inclusion of students whose scores on the lowest-rated dimension, identifying bias and assumptions, varied, so that the dimension most in need of explanation was well covered; and capacity to provide articulate, detailed accounts of their reading, judged from open responses and willingness to participate. Selecting on the quantitative scores in this way is what made the qualitative phase explanatory rather than merely sequential, because it directed

the interviews toward the patterns that the numbers had raised. Participants are identified by the codes P1 to P20 to preserve confidentiality.

Instruments

Two instruments were used. The questionnaire comprised 10 statements (CR1–CR10) rated on a five-point Likert scale from 1 (strongly disagree) to 5 (strongly agree), with two items for each of the five dimensions. The items were developed by operationalizing the critical reading construct described above (Wallace, 2003; Mulumba, 2016) and the source-evaluation emphasis of critical digital literacy work (Silvhiany & Huzaifah, 2021), each dimension being defined before items were written. Identifying main ideas and arguments was defined as locating the central claim and the author's line of reasoning; analyzing information as tracing relationships among ideas and separating claims from supporting evidence; evaluating source credibility as judging the trustworthiness of authors and sources; identifying bias and assumptions as detecting one-sidedness and unstated premises; and drawing conclusions and inferences as synthesizing evidence into warranted judgments. Two sample items illustrate the format: "I can identify the author's main argument in an online text" (CR2) and "I check the credibility of the sources of information I read online" (CR5).

The semi-structured interview protocol contained open-ended questions aligned one-to-one with the five dimensions, each asking students to describe the strategies and reasoning behind their reading. Because every interview question mapped onto a measured dimension, the protocol was positioned to explain the questionnaire results directly rather than to open unrelated lines of inquiry.

Validity, Reliability, and Trustworthiness

Content validity of the questionnaire was established through expert review. Three lecturers in English education and educational measurement independently rated the relevance of each item to its intended dimension; item-level content validity indices ranged from .80 to 1.00 and the scale-level content validity index (averaging method) was .92, indicating satisfactory content representation. Internal consistency, computed on the 20 responses, was good, with a Cronbach's alpha of .86 for the full 10-item scale, and corrected item-total correlations were all positive and acceptable; these values are reported as descriptive evidence of measurement adequacy for a short, dimension-aligned instrument rather than as a substitute for full-scale validation, which the limited sample did not permit.

For the qualitative phase, trustworthiness was pursued against the four criteria of credibility, transferability, dependability, and confirmability (Lincoln & Guba, 1985; Nowell et al., 2017). Credibility was supported by triangulating questionnaire and interview data and by member checking, in which selected participants confirmed that summarized meanings reflected their intentions. Transferability was supported by thick description of the setting and participants. Dependability and confirmability were supported by an audit trail documenting coding decisions, by peer debriefing between the two coders, and by returning to the transcripts whenever interpretation was uncertain. Before administration, items and interview questions were reviewed for clarity and fit, and minor wording was adjusted to remove ambiguity.

Researcher Reflexivity

Because the qualitative data were generated and interpreted by the researchers, their position warrants acknowledgement. The researchers are teacher-educators within the same English Education program, which gave them contextual familiarity and ready access but also an insider relationship to participants that could encourage socially desirable responses or shared assumptions about what good reading looks like. To limit this influence, interviews emphasized concrete descriptions of what students actually did rather than evaluations of whether they read well, a reflexive journal was kept to record the researchers' reactions and provisional interpretations, and coding was conducted independently by two researchers before reconciliation, so that no single perspective silently shaped the themes.

Data Analysis and Integration

Quantitative data were analyzed using descriptive statistics, namely the mean and standard deviation for each item and dimension and an overall composite mean. Mean scores were interpreted with a three-band scale derived from the response range: 1.00–2.33 low, 2.34–3.67 moderate, and 3.68–5.00 high. Qualitative data were analyzed through the six phases of reflexive thematic analysis (Braun & Clarke, 2006; Nowell et al., 2017): familiarization, generating initial codes, searching for themes, reviewing themes, defining and naming themes, and producing the report. Coding stayed close to the strategies students described; 38 initial codes were generated, then grouped into six subthemes and finally into two overarching themes. As an example of the coding chain, the interview segment "I compare it with other sources and check who wrote

it" was coded as compare across sources and check author, which fed the subtheme cross-source credibility checking and the theme evaluative and reflective reading practices. The full coding structure is reported in Table 1.

Integration was conducted as a planned analytic step rather than a reporting convenience. A joint display was constructed by aligning each quantitative dimension with the qualitative theme that addressed it and recording an integrated interpretation (Fetters et al., 2013; Guetterman et al., 2015). Each pairing was then assigned a fit judgment, following the logic of confirmation, expansion, or discordance, and meta-inferences were drawn from the pattern of fits. The quality of the integration was appraised against mixed methods legitimation criteria, attending to whether the inferences from the two strands were consistent and whether their combination was warranted by the data (Onwuegbuzie & Johnson, 2006; O'Cathain et al., 2008).

Table 1. Coding Structure: Themes, Subthemes, and Representative Codes

Theme	Subtheme	Representative codes (illustrative)
Theme 1: Strategic critical engagement with digital texts	1a Orienting to text structure	title/abstract/conclusion first; keyword scanning
	1b Tracing argument and idea relationships	link ideas before judging; read several articles on a topic
	1c Distinguishing fact from opinion	check for data/references; no evidence = opinion
Theme 2: Evaluative and reflective reading practices	2a Cross-source credibility checking	compare across sources; check author and date; site reputation
	2b Comparative bias detection	read several sites on an issue; notice leaning/positioning
	2c Evidence-based synthesis	conclude from evidence; not first impression

Note. Thirty-eight initial codes were collapsed into the six subthemes and two themes shown; codes listed are representative rather than exhaustive.

Results and Discussions

The results are reported in three parts, reflecting the explanatory sequential logic. The quantitative findings establish the level of students' critical reading; the qualitative findings, organized as two themes with six subthemes, explain that level; and the integrated findings merge the two through a joint display and the meta-inferences it supports.

Quantitative Findings

Across the 10 items, students rated their critical reading through digital texts highly, with an overall composite mean of 4.52 (SD = 0.50), which falls within the high band of the interpretive scale. Item-level results, reported in Table 2, ranged from 4.40 to 4.65, indicating that students rated themselves consistently across the individual statements rather than strongly on some and weakly on others.

Table 2. Item-Level Descriptive Statistics (N = 20)

Item	Statement (abridged)	Mean	SD
CR1	Identifies the main idea in digital texts	4.50	0.51
CR2	Identifies the author's main argument	4.60	0.50

Item	Statement (abridged)	Mean	SD
CR3	Analyzes relationships among ideas	4.50	0.51
CR4	Distinguishes facts from opinions	4.40	0.50
CR5	Checks credibility of information sources	4.55	0.51
CR6	Considers author or website reputation	4.55	0.51
CR7	Recognizes bias in digital texts	4.45	0.51
CR8	Identifies assumptions behind arguments	4.40	0.50
CR9	Draws conclusions from digital texts	4.65	0.49
CR10	Makes logical inferences from digital texts	4.60	0.50

Note. Responses were rated from 1 (strongly disagree) to 5 (strongly agree).

Aggregated to the five dimensions, the same pattern of consistently high attainment held. Drawing conclusions and inferences received the highest dimension mean, followed by identifying main ideas and arguments and evaluating source credibility, then analyzing information, with identifying bias and assumptions the lowest, though still in the high band. Table 3 and Figure 3 summarize the dimension-level results.

Table 3. Dimension-Level Descriptive Statistics and Category (N = 20)

Dimension	Items	Mean	SD	Category
Identifying main ideas and arguments	CR1–CR2	4.55	0.50	High
Analyzing information	CR3–CR4	4.45	0.50	High
Evaluating source credibility	CR5–CR6	4.55	0.50	High
Identifying bias and assumptions	CR7–CR8	4.42	0.50	High
Drawing conclusions and inferences	CR9–CR10	4.62	0.49	High
Total	**CR1–CR10**	**4.52**	**0.50**	**High**

Note. Category interpretation: 1.00–2.33 = low; 2.34–3.67 = moderate; 3.68–5.00 = high. Cronbach's α = .86.

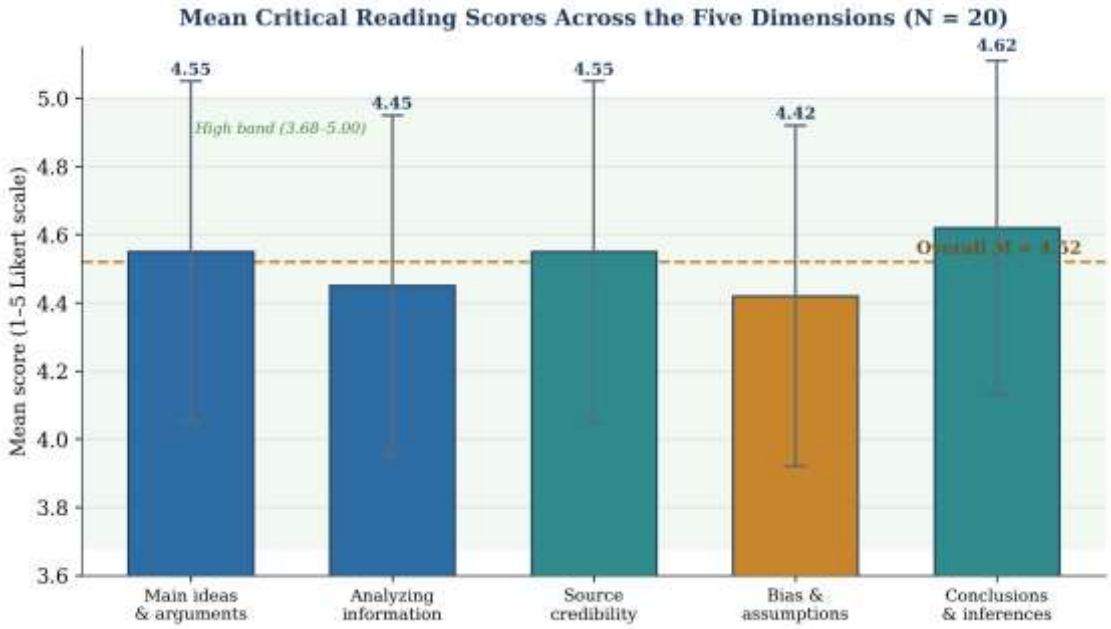

Figure 3. Consistently High Scores Across Dimensions

Read on their own terms, these numbers establish that students perceived themselves as competent critical readers across the full range of operations measured. Two features of the distribution, however, call for caution rather than celebration. First, the scores are uniformly high and tightly clustered, with standard deviations near half a scale point, a pattern that can reflect genuine competence but can equally reflect self-report and social-desirability bias, since respondents may overstate skills they know to be valued. Second, self-rated questionnaires capture perceived rather than demonstrated ability. For both reasons the quantitative phase is treated here as establishing a level to be explained and probed, not as a finished claim about competence; the qualitative phase was designed precisely to test whether behaviorally specific accounts would corroborate or undercut the scores.

Qualitative Findings

Thematic analysis of the 10 interviews produced two themes spanning six subthemes, mapped in Figure 4. The first theme, strategic critical engagement with digital texts, captures the analytical practices through which students worked with a text to understand it; the second, evaluative and reflective reading practices, captures the judgments they made about trustworthiness and the way they reasoned toward conclusions. Together the themes account for the high scores, but the subtheme structure also exposes where that account is uneven.

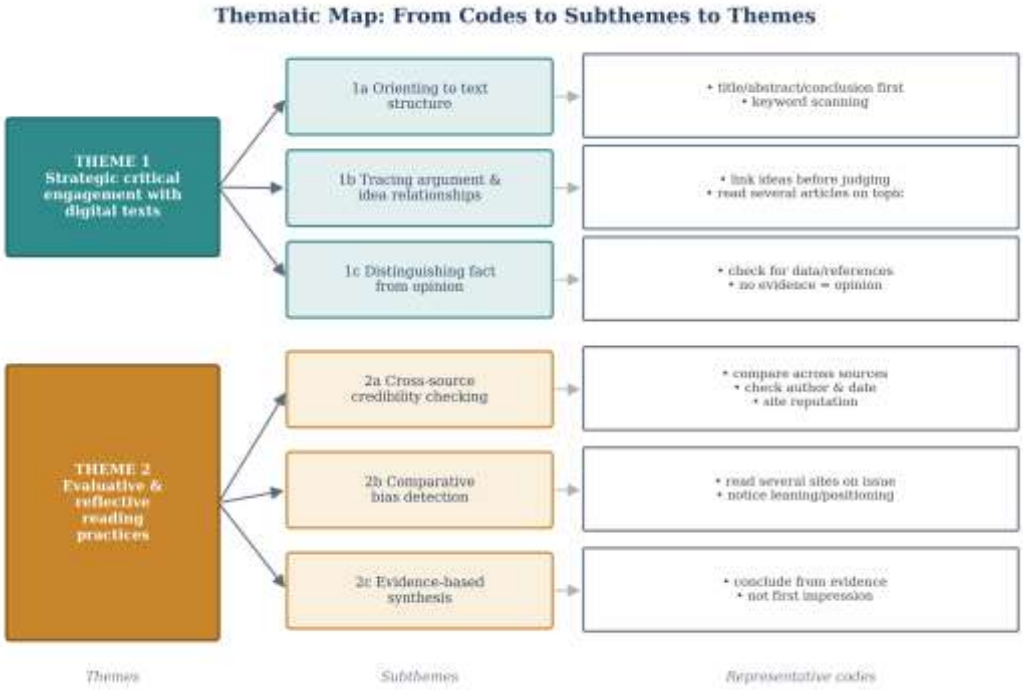

Figure 4. Thematic Map of Reading Strategies

Theme 1: Strategic Critical Engagement with Digital Texts.

Students described approaching digital texts as active analysts rather than passive recipients. Their first move was typically to orient to the structure of a text in order to locate its main idea before reading in detail (subtheme 1a).

> "Saya biasanya mulai dari judul, abstrak, dan kesimpulan, karena itu membantu saya menangkap ide utama sebelum membaca keseluruhan teks secara rinci." ("I usually begin with the title, abstract, and conclusion, because they help me capture the central idea before I read the whole text in detail.") (P1)

> "Saya mencari kata kunci dan ide pokok dulu, supaya saya tahu sebenarnya teks itu membahas apa." ("I look for keywords and the main idea first, since that lets me see what the whole text is really about.") (P12)

Having located the central point, students traced how ideas related and weighed whether statements were supported by evidence (subtheme 1b), and distinguished facts from opinions by the presence or absence of data and references (subtheme 1c).

> "Saya berusaha melihat bagaimana gagasan-gagasan dalam teks saling berhubungan sebelum memutuskan apa yang sebenarnya penulis argumentasikan." ("I try to see how the ideas in the text connect to one another before I decide what the writer is actually arguing.") (P7)

> "Untuk memisahkan fakta dari opini, saya cek apakah penulis memberi data atau referensi; kalau tidak ada bukti, saya anggap itu opini penulis." ("To separate facts from opinions, I check whether the writer gives data or references; if there is no evidence, I treat it as the author's opinion.") (P4)

These accounts explain why the dimensions of identifying main ideas and arguments and analyzing information were rated highly: students held routine, deliberate procedures for finding a text's central point and examining its internal logic, and applied them fluently to digital material. Not every account was uniform, however. One student admitted relying on familiarity rather than analysis when pressed for time, a divergence that tempers the otherwise consistent picture.

> "Jujur, kalau sedang buru-buru, kadang saya langsung percaya kalau sumbernya sudah terkenal, tanpa menganalisis isinya lebih dalam." ("Honestly, when I am in a hurry, sometimes I just trust a source if it is already well known, without analyzing the content more deeply.") (P9)

Theme 2: Evaluative and Reflective Reading Practices.

The second theme concerns the judgments students made before accepting information and the way they reasoned toward conclusions. Evaluating credibility was the most frequently mentioned activity: students cross-checked information and inspected the author and source (subtheme 2a).

> "Saya membandingkan informasi dari beberapa sumber digital untuk memastikan bahwa apa yang saya baca bisa dipercaya." ("I compare information across several digital sources to make sure that what I read can be trusted.") (P2)

> "Sebelum percaya pada informasi daring, saya selalu membandingkannya dengan sumber lain dan mengecek kredibilitas penulisnya." ("Before trusting online information, I always compare it with other sources and check the author's credibility.") (P11)

Bias detection followed the same comparative logic (subtheme 2b). Rather than dissecting a single text from within, students read several sources on an issue and noticed where one leaned, a strategy that is effective but external to the text being judged.

> "Saya membaca beberapa situs tentang isu yang sama untuk melihat apakah suatu teks cenderung memihak satu sisi." ("I read several websites on the same issue to notice whether a particular text is leaning toward one side.") (P8)

Finally, students drew conclusions by synthesizing evidence rather than reacting to first impressions (subtheme 2c).

> "Saya menarik kesimpulan dari bukti yang tersedia dalam teks, bukan hanya dari kesan pertama." ("I draw my conclusions from the evidence available in the text, not just from my first impression.") (P15)

This theme explains the high scores for evaluating source credibility, identifying bias and assumptions, and drawing conclusions and inferences. It also explains a subtler pattern: because bias was detected by comparison across texts rather than by analyzing the rhetoric and assumptions within one text, that dimension, while still high, was the lowest, and the strategy behind it is partial. Figure 5 displays the concepts most frequently coded across the transcripts. The word cloud indexes salience rather than relationships, so it is reported as a descriptive complement to the thematic map in Figure 4 rather than as analysis in its own right; terms such as source credibility, compare sources, cross-check, and draw conclusions recur most strongly, consistent with the two themes.

Figure 5. Most Frequent Concepts in Interviews

Integrated Findings

The joint display in Table 4 connects the quantitative dimensions to the two themes, assigns a fit judgment to each pairing, and records the resulting interpretation; Figure 6 visualizes the same logic. On four of the five dimensions the strands stood in a relationship of confirmation: the strategies students described provided a plausible and behaviorally specific account of the high scores, and the convergence strengthens confidence that the measured level reflected genuine practice rather than self-report alone. On the dimension of identifying bias and assumptions the relationship was one of expansion rather than simple confirmation, because the interviews supplied information the questionnaire could not: students scored this dimension high yet detected bias externally, by cross-source comparison, rather than by interrogating a single text. This is the analytic payoff of integration, since it is precisely the kind of insight neither strand could have produced alone.

Table 4. Joint Display Linking Quantitative Findings, Qualitative Themes, Fit, and Integrated Interpretation

Quantitative findings	Qualitative theme/subthemes	Fit	Integrated interpretation (meta-inference)
Main ideas & arguments (M = 4.55); Analyzing information (M = 4.45)	Theme 1: Strategic critical engagement (1a–1c)	Confirmation	Fluent analytic routines for locating ideas and checking evidence account for the high analytic scores.
Source credibility (M = 4.55); Conclusions & inferences (M = 4.62)	Theme 2: Evaluative & reflective practices (2a, 2c)	Confirmation	Cross-source verification and evidence-based synthesis explain the strong appraisal and inference scores.
Bias & assumptions (M = 4.42, lowest)	Theme 2: Comparative bias detection (2b)	Expansion	High self-rating rests on an external, across-text strategy rather than within-text analysis; the qualitative data add why this dimension is weakest and mark a developmental edge.

Note. Fit categories follow the logic of confirmation, expansion, and discordance. No discordance was observed, but the expansion on bias shows the strands did more than echo each other.

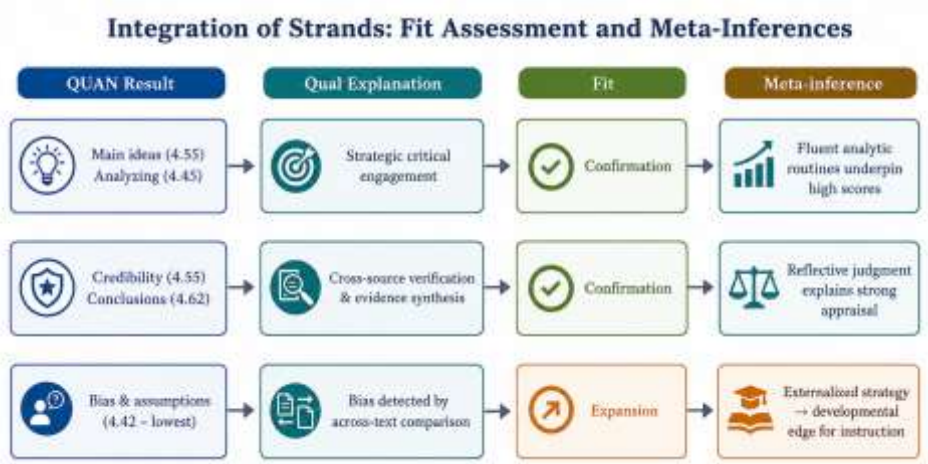

Figure 6. Fit Assessment and Meta-Inference Integration

Taken together, the integrated findings indicate that the high level measured in the first phase rested on a coherent and largely shared repertoire of reading strategies documented in the second, while the one expansion case tempers the uniformly rosy quantitative picture by showing that not all of the high scores rest on equally deep strategies. The meta-inference is therefore twofold: students' self-rated competence is broadly corroborated by their described practice, and the dimension they rate lowest is also the one whose underlying strategy is most external to the text, which is why it is both the weakest score and the clearest target for instruction.

Set against the wider literature, these results speak back to the inconsistency identified in the introduction. The strength of source evaluation here contrasts with studies that found EFL learners weak at judging online sources (Silvhiany et al., 2021) and aligns instead with the verification routines that define critical digital literacy as conceptualized and assessed in recent work (Silvhiany & Huzaifah, 2021; Ismael et al., 2026). The most plausible reading is not that one body of findings is wrong but that the competence is contingent: these participants were purposively engaged readers of academic digital texts in a program where such reading is routine, and teachers' practices are known to shape whether the skill develops (Rini & Nabhan, 2023). The fluent structural routines students reported resemble the strategic behaviors that distinguish skilled digital readers in comparative work (Al-Amrani, 2022; Rustambekova, n.d.) and support the view that digital environments scaffold reading when engagement is habitual (Asia et al., 2024; Mistang et al., 2025; Ramadhini & Wahyuni, 2026). The pattern is consistent, too, with intervention and narrative studies showing that structured or responsive pedagogy can build critical reading in digital settings (Hidayanto et al., 2026; Iswati et al., 2025).

Within critical reading more narrowly, the two themes map onto the cultivable operations that define the construct (Wallace, 2003; Mulumba, 2016; Litim & Bouguebs, 2025): strategic engagement corresponds to locating and analyzing, evaluative reflection to appraising, detecting bias, and inferring. That conclusions and inferences scored highest, together with students' descriptions of evidence-based synthesis, suggests these readers did not stop at surface comprehension, which is encouraging given concerns that the digital environment fragments attention (Khusniyah, 2024), and the fact that students themselves regard critical reading as demanding (Julianti et al., 2024) makes the reported competence more, not less, meaningful. The expansion finding on bias gives the discussion its sharpest point: detecting bias by comparing sources is effective but partial, and within-text interrogation of assumptions and positioning is the skill instruction can most usefully add, through collaborative, critically designed, and multimodal materials (Hernandez-Caicedo, 2023; Puspitasari et al., 2025; Wafiroh et al., 2024; Pakarasang, 2025; Donkol, 2023). This echoes evidence that critical thinking and digital literacy do not advance in step among Indonesian EFL students (Indah et al., 2022).

Two cautions qualify the interpretation. The convergence between strands raises confidence but does not eliminate the self-report limitation, since interviews are also self-described practice rather than observed performance; the single divergent account (P9) is a reminder that the homogeneity of the scores may smooth over real variation. And although the study speaks of digital texts contributing to critical reading, its cross-sectional, non-comparative design cannot establish such a causal contribution; digital texts are best read here as the affordance within which these strategies were exercised, consistent with the framing in Figure 1.

Conclusions

The integrative contribution of this study is best stated as a meta-inference rather than a pair of separate results: among these Indonesian EFL undergraduates, a uniformly high self-rated level of critical reading was corroborated, dimension by dimension, by concrete strategic and evaluative practices, with one telling exception. The exception is the heart of the finding. On identifying bias and assumptions, the lowest-rated dimension, the qualitative evidence expanded rather than merely confirmed the quantitative result, showing that students sustained a high rating through an external strategy, comparing across texts, rather than through within-text analysis. Neither strand alone could have produced this understanding: the questionnaire would have left a uniformly high profile unexplained, and the interviews would have described strategies without revealing which dimension they explained least well.

This is also where the qualitative phase modified, and did not simply echo, the quantitative interpretation. By exposing the externalized basis of bias detection and surfacing a divergent case, the interviews tempered

the impression of uniform mastery that the numbers alone conveyed, and they redirected attention from a general claim of competence toward a specific developmental edge. Read in this light, the study's contribution is twofold. Theoretically, it offers a worked example of how an explanatory sequential design can connect measured competence to documented practice in critical reading research, and of how a fit assessment within a joint display can yield meta-inferences beyond either strand (Fetters et al., 2013; Guetterman et al., 2015; Marboot et al., 2020). Methodologically, the deliberate alignment of questionnaire items, interview questions, and joint-display rows offers a transferable template for making critical reading both measurable and explicable.

The conclusions are bounded by the design. The sample was small, drawn from a single program, and composed of engaged digital readers, so the high level should not be generalized. Both strands rest on self-report rather than observed performance, and the cross-sectional design measures an attained state at one time point rather than change, which is why claims about development and about the contribution of digital texts are framed here as description and affordance rather than as growth or cause. These limits point directly to the work the integration makes most pressing.

For pedagogy, the integrated evidence yields a specific rather than general recommendation: build on the source-evaluation and synthesis strengths students already demonstrate, and teach explicitly the within-text analysis of bias and assumptions that the integration identified as their weakest underlying strategy. For research, the same integration sets the agenda. Studies that actually measure change over time would let the word development carry its full meaning; performance-based and think-aloud measures would test whether self-described strategies survive contact with real texts; larger and more varied samples would establish how far the pattern holds; and intervention studies targeting within-text bias detection would test whether the developmental edge identified here can be moved. More broadly, in an EFL setting where digital texts are a primary source of authentic English, helping students read those texts critically remains inseparable from helping them learn the language itself.

References

Afnan, A. (2023). Digital literacies for English language learners: A multimodal social semiotic approach to investigate Saudi students' reading of English tweets [Doctoral dissertation, Newcastle University].

Al-Amrani, S. N. (2022). Learning strategies used by EFL learners for print and digital reading: Critical review. SSRN. https://doi.org/10.2139/ssrn.4000436

Asia, M., Asdam, M., & Asdar, A. (2024). Investigating students' reading strategies and reading comprehension through digital literacy environment. AL-ISHLAH: Jurnal Pendidikan, 16(3), 3370–3379.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Bryman, A. (2006). Integrating quantitative and qualitative research: How is it done? Qualitative Research, 6(1), 97–113. https://doi.org/10.1177/1468794106058877

Coiro, J. (2011). Predicting reading comprehension on the Internet: Contributions of offline reading skills, online reading skills, and prior knowledge. Journal of Literacy Research, 43(4), 352–392. https://doi.org/10.1177/1086296X11421979

Creswell, J. W., & Plano Clark, V. L. (2018). Designing and conducting mixed methods research (3rd ed.). SAGE Publications.

Donkol, S. A. A. S. M. (2023). Using an electronic interactive program for developing EFL analytical writing and critical reading skills of secondary stage students [Doctoral dissertation, Minia University].

Fetters, M. D., Curry, L. A., & Creswell, J. W. (2013). Achieving integration in mixed methods designs—Principles and practices. Health Services Research, 48(6pt2), 2134–2156. https://doi.org/10.1111/1475-6773.12117

Greene, J. C. (2007). Mixed methods in social inquiry. Jossey-Bass.

Guetterman, T. C., Fetters, M. D., & Creswell, J. W. (2015). Integrating quantitative and qualitative results in health science mixed methods research through joint displays. Annals of Family Medicine, 13(6), 554–561. https://doi.org/10.1370/afm.1865

Hernandez-Caicedo, A. L. (2023). Critical reading through decolonized print teacher-designed ELT materials otherwise.

Hidayanto, H. N., Sumardjoko, B., & Sutopo, A. (2026). Innovating critical reading through HOTS-based digital books: A mixed-methods study of the FIVES strategy among EFL university students. Jurnal Cakrawala Pendidikan, 45(1).

Indah, R. N., Toyyibah, T., Budhiningrum, A. S., & Afifi, N. (2022). The research competence, critical thinking skills and digital literacy of Indonesian EFL students. Journal of Language Teaching and Research, 13(2), 315–324.

Ismael, I., Li, S., & Luo, X. (2026). The 21st century EFL literacy at stake: Extending language literacy to critical and digital empowerment. Asian Journal of Education and Social Studies, 52(4), 374–382.

Iswati, H., Munir, A., & Nugroho, H. M. A. (2025). Dynamic pedagogy for critical thinking: A narrative inquiry in digital EFL reading. International Journal of Pedagogical Language, Literature, and Cultural Studies (i-Plural), 2(3), 175–191.

Javahery, P., Alpat, M. F., & Kamali, J. (2026). #LearnEnglish: Social media, language learning, and the need for critical digital literacy. Journal of Applied Learning & Teaching, 9(1), 1–20.

Julianti, R., Anggung, S. D., & Ima, N. (2024). Undergraduate EFL students' perspective on critical reading. ELS Journal on Interdisciplinary Studies in Humanities, 7(2), 376–382.

Khusniyah, N. L. (2024). Critical reading for critical thinking: The urgency of critical reading in the era of technological and information disruption. Proceeding of International Conference on Islamic Education and Science Development, 2(1), 183–189.

Leu, D. J., Kinzer, C. K., Coiro, J., Castek, J., & Henry, L. A. (2013). New literacies: A dual-level theory of the changing nature of literacy, instruction, and assessment. In D. E. Alvermann, N. J. Unrau, & R. B. Ruddell (Eds.), Theoretical models and processes of reading (6th ed., pp. 1150–1181). International Reading Association.

Lincoln, Y. S., & Guba, E. G. (1985). Naturalistic inquiry. SAGE Publications.

Litim, H., & Bouguebs, R. (2025). The role of critical literacy in enhancing EFL reading skills: A case study of fourth-year graduating EFL learners at ENS Constantine. ALTRALANG Journal, 7(2), 175–187.

Luthfia, A. F. (2025). Digital literacy for developing reading skills in EFL contexts: A systematic review. Aplinesia (Journal of Applied Linguistics Indonesia), 9(2), 40–64.

Marboot, K., Roohani, A., & Mirzaei, A. (2020). Investigating Iranian EFL students' metacognitive online reading strategies, critical thinking, and their relationship: A mixed-methods study. Issues in Language Teaching, 9(1), 151–182.

Mistang, M., Ruslin, R., Rengur, Z. A., Syam, H., & Wirana, S. (2025). Digital media in teaching & learning reading comprehension: The view of the EFL pre-service teachers in an Indonesian higher education. Jurnal Pendidikan: Teori, Penelitian, dan Pengembangan, 10(8), 343–351.

Mulumba, M. B. (2016). Improving student teachers' knowledge-base in language education through critical reading. Pedagogy, Culture & Society, 24(4), 571–586.

Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic analysis: Striving to meet the trustworthiness criteria. International Journal of Qualitative Methods, 16(1), 1–13. https://doi.org/10.1177/1609406917733847

O'Cathain, A., Murphy, E., & Nicholl, J. (2008). The quality of mixed methods studies in health services research. Journal of Health Services Research & Policy, 13(2), 92–98. https://doi.org/10.1258/jhsrp.2007.007074

Onwuegbuzie, A. J., & Johnson, R. B. (2006). The validity issue in mixed research. Research in the Schools, 13(1), 48–63.

Pakarasang, C. (2025). Guidelines for designing EFL textbooks to promote critical reading using out-of-class English materials.

Puspitasari, A., Jailani, M. K., Basthomi, Y., Suharyadi, S., & Zen, E. L. (2025). Multimodal scaffolding in critical reading: Pedagogical insights from Indonesian EFL learners. Celtic: A Journal of Culture, English Language Teaching, Literature and Linguistics, 12(2), 924–947.

Rahayu, Y. S., Mujiyanto, J., & Suwandi, S. (2023). The impact of electronic storybook and print picture book on the reading comprehension of undergraduates with varying levels of critical thinking. Migration Letters, 20(S1), 224–237.

Ramadhini, N. A., & Wahyuni, I. F. (2026). Exploring EFL students' perceptions of online articles for critical reading. Academia Open, 11(1), 10–21070.

Rini, H. M., & Nabhan, S. (2023). Exploring critical digital literacy in a higher education EFL classroom: Teachers' views and practices. ELT Forum: Journal of English Language Teaching, 12(3), 216–228.

Rustambekova, O. (n.d.). Digital reading vs. print reading: Effects on EFL learners' reading speed and understanding. Maktabgacha va Maktab Ta'limi Jurnali.

Silvhiany, S., & Huzaifah, S. (2021, January). Developing online critical literacy assessment for EFL learners. In 4th Sriwijaya University Learning and Education International Conference (SULE-IC 2020) (pp. 378–386). Atlantis Press.

Silvhiany, S., Huzaifah, S., & Ismet, I. (2021). Critical digital literacy: EFL students' ability to evaluate online sources. Indonesian Journal of EFL and Linguistics, 6(1), 249.

Teddlie, C., & Tashakkori, A. (2009). Foundations of mixed methods research: Integrating quantitative and qualitative approaches in the social and behavioral sciences. SAGE Publications.

Wallace, C. (2003). Critical reading in language education. Palgrave Macmillan.

Wafiroh, L. H., Retnaningdyah, P., & Munir, A. (2024). EFL students' critical reading and collaborative learning skills via virtual literature circle (VLC) within a critical reading class. Applied Research on English Language, 13(3), 65–90.