ISSN 1799-2591 (print); ISSN 2053-0692 (online)
Theory and Practice in Language Studies, Vol. 16, No. 11, pp. 3997-4007, November 2026
DOI: https://doi.org/10.17507/tpls.1611.31

Comprehension Breakdown and Regulatory Strategies in Academic Reading: A Think-Aloud Case Study of University Students in Kupang, Indonesia

Lenny Nofriyani Adam
Linguistics Study Program, Universitas Pendidikan Indonesia, Bandung, Indonesia;
Universitas Timor, Indonesia

Dadang Sudana*
Linguistics Study Program, Universitas Pendidikan Indonesia, Bandung, Indonesia

Heru Pratikno
Linguistics Study Program, Universitas Pendidikan Indonesia, Bandung, Indonesia

Joni Soleman Nalenan
Linguistics Study Program, Universitas Pendidikan Indonesia, Bandung, Indonesia

Yandris Obligus Nabuasa
General and Character Education, Universitas Pendidikan Indonesia, Bandung, Indonesia

Hanny Withney Gloria Wayangkau
English Education Study Program, Universitas Pendidikan Indonesia, Bandung, Indonesia

Fauzan Novaldy Pratama
Linguistics Study Program, Universitas Pendidikan Indonesia, Bandung, Indonesia

***Abstract*—Academic reading is central to the performance of higher education students, but how students manage difficulties in reading comprehension remains under-researched. This study examined the comprehension breakdown and regulatory strategies that students verbalized during reading, focusing on strategies readers use to restore understanding and increase comprehension. Nine second-semester university students in Kupang, Indonesia, were involved in a think-aloud reading task followed by stimulated recall interviews. The results showed that academic reading unfolded in several episodes rather than as a smooth linear process when they found difficult technical terms, abbreviations, abstract conceptual labels, and underspecified evidential claims. Three main strategies to achieve comprehension were repair chains, concretization bridges to recover abstract meaning, and the application of meaning monitoring or epistemic monitoring. Across cases, participants varied in the density and depth of regulatory strategies; some produced longer, less stable episodes, while others often recovered meaning through glossing, paraphrasing, inference, and evaluative monitoring.**

***Index Terms*—academic reading, comprehension breakdowns, think-aloud protocol, Kupang students**

I. Introduction

Reading activity involves more than decoding written symbols (Hung et al., 2025; Paige et al., 2024)—it takes the capability of constructing meaning, evaluating textual coherence, and making interpretive judgments (Alruthaya et al., 2025; Kiili et al., 2024; McCrudden et al., 2024; Tibken & Tiffin-Richards, 2025). Recent scholarship reports that reading comprehension does not emerge from isolated subskills alone, but from the dynamic interactions between the reader's characteristics, text features, and context (Duke & Cartwright, 2021; McMaster & Kendeou, 2023). Today, the international assessment data continue to indicate persistent challenges in reading achievement. For example, the 2022 PISA report stated that 15-year-old Indonesian teenagers generally scored 359 in reading comprehension, with only 25% attained Level 2 and above, compared to the OECD average of 476 (OECD, 2023). While this nationwide finding indicates a persistent issue in literacy, it lacks information on how comprehension difficulties arise during the reading process.

In higher education, students are required to digest academic texts with more conceptual density, specialized lexicon, and complex structures. Beyond literal meaning, students must be able to connect propositions across clauses and

© 2026 ACADEMY PUBLICATION

paragraphs, relate textual information to their prior knowledge, and construct a coherent situation model of the text (Adams, 1990; Chall, 1983; Kintsch, 1998; Perfetti, 1985; Rayner et al., 2012; Smith, 2004; Snow & Sweet, 2003). As such, recent studies suggest that students often perceive reading comprehension—the key factor of academic participation and success—as a demanding and frustrating skill (Bergman, 2024; Mason & Warmington, 2024). In general, research on reading comprehension of higher education students remains outcome-oriented, focusing on comprehension scores, grades, or readers' self-reported strategies rather than the real-time processes through which understanding breaks down and is subsequently repaired. Some works from different parts of the world also reported that most students find it challenging to develop academic literacies (Lea & Street, 1998), which significantly influence their success in higher education (Atkinson & Smith Risser, 2023; Holschuh, 2020; Reynolds et al., 2020).

Successful reading depends not only on cognitive skills (linguistic processing) but also on metacognitive monitoring and metacognitive regulation. Cognitive skills involve mental processes required to attend to, interpret, store, and connect information in a text. More specifically, linguistic processing allows readers to comprehend vocabulary, parse syntactic structures, and establish semantic relationships within and across sentences. Metacognitive monitoring refers to a reader's ability to observe and evaluate ongoing comprehension, particularly in identifying misunderstanding, uncertainty, or gaps in interpretation. Metacognitive regulation is the deliberate actions taken to manage and repair comprehension, such as rereading difficult passages, inferring meaning from context, or adjusting reading strategies to task demands.

Readers experience what is called comprehension breakdown when they encounter difficulties in reading, such as unfamiliar technical terms, abstract concepts, or weakly signaled relations between ideas in the texts. To address this, they may identify the source of difficulty and try to increase comprehension by rereading, paraphrasing, inferring, or summarizing (Flavell, 1979; Nelson & Narens, 1990). Recent research continues to show that text with conceptually difficult content can disrupt readers' accurate judgment, so metacognitive monitoring is a useful strategy to increase comprehension (Hildenbrand et al., 2024; Tibken & Tiffin-Richards, 2025). Accordingly, research on reading comprehension should distinguish between comprehension breakdown, metacognitive monitoring, and regulatory strategies to identify how readers notice the difficulties and attempt to resolve them.

A substantial body of research mostly uses declarative instruments, such as strategy questionnaires or retrospective reports (Craig et al., 2020; Schellings & Van Hout-Wolters, 2011), to identify general tendencies, but it is less effective to capture online processing and vulnerable to recall bias and social desirability (Paulhus, 2017; Veenman, 2011). Recent scholarship has called for a closer alignment between theories of reading comprehension and the assessment and investigation of comprehension, particularly if the field is to account for the dynamic interaction between the reader, the text, and the context (McMaster & Kendeou, 2023). To understand how readers experience difficulty and attempt to recover from it, a more process-sensitive approach is needed.

This study addresses the need through the think-aloud protocol (TAP). Think-aloud is a research method in which participants speak aloud any words in their mind as they complete a task (Charters, 2003). Within the protocol-analysis tradition, think-aloud data provide structured access to verbalizable aspects of cognitive processing during task performance (Ericsson & Simon, 1993; Pressley & Afflerbach, 1995). Tang et al. (2025) suggest that think-aloud remains a valuable method for examining reading comprehension, strategic behavior, and metacognitive activity, particularly in language-related contexts. To mitigate the limitations associated with relying solely on verbal protocol data and to strengthen the credibility of the findings, this study triangulated TAP data with stimulated recall interviews (SRI) conducted after the reading task. This allowed participants to elaborate on moments of hesitation, uncertainty, or strategic decision-making, thereby supporting a more comprehensive interpretation of their cognitive and metacognitive processes during academic reading (Fan et al., 2023).

It has been reported that a careful use of stimulated recall alone, or combined with other techniques, remains a productive way to access reflections on cognitive and affective processes in educational research (Zhai et al., 2024). The combination of TAP and SRI has been used in various research contexts to reveal participants' cognitive processes. Bowles (2018) explains that think-aloud and stimulated recall are forms of verbal reports that have long been used in language research to access learners' thinking processes. Empirically, this combination has been applied in the contexts of sport coaching, EFL writing, feedback-based revision, and EFL writing processes in Indonesia (Abas & Abd Aziz, 2018; Chapman et al., 2024; Hoang & Storch, 2024; Tian et al., 2022). Even in the Indonesian context, Supardan and Capati (2026) employed academic reading tasks, screen recordings, think-aloud protocols, AI interaction logs, and stimulated recall interviews to examine how EFL students use AI as a metacognitive mirror to monitor and repair comprehension breakdowns. However, studies that specifically integrate TAP and SRI to uncover students' cognitive processes in reading academic expository texts without AI mediation, particularly in the context of a university in Kupang City, remain limited. Therefore, this study focuses on students' cognitive processes as they engage directly with academic expository texts.

This research was conducted in Kupang, the capital city of East Nusa Tenggara Province, Indonesia. This area is selected because it receives less attention in previous TAP and SRI-based studies and is underrepresented in the international literature of academic reading. This study examines the cognitive processes that are verbalized during reading, focusing on the strategies readers use to restore understanding and achieve comprehension. Case-based evidence from such contexts can enrich broader theoretical discussions by showing how comprehension breakdown, monitoring, and regulation unfold in settings that remain underexamined.

© 2026 ACADEMY PUBLICATION

II. Methodology

This study employed a qualitative case study to explore the verbalized cognitive processes involved in academic reading. The participants were nine second-semester students of the Indonesian Language Department in a university in Kupang, who enrolled in the Indonesian Basic Reading Course. These participants were purposively selected based on their levels of UKBI (a standardized assessment of Indonesian proficiency): Semenjana (basic), Madya (intermediate), and Unggul (fluent), with three students from each level. The UKBI level only served as an interpretive background rather than a basis for statistical comparison. The participants were coded with S1-S3 (Semenjana), M1-M3 (Madya), and U1-U3 (Unggul). In the preliminary stage, the participants were asked to fill out a questionnaire about their language use, reading exposure, prior knowledge, and topic interest.

The primary data were the verbatim transcription of the think-aloud protocol (TAP) activities combined with the stimulated recall interview (SRI) as the supporting evidence. In the TAP sessions, the participants were asked to read an expository text about the application of technology in higher education. The 500-600-word text was in Bahasa Indonesia, divided into eight paragraphs. It contained technical terms, abstract concepts, and generalized claims to stimulate monitoring and regulation. While reading, the participants should say whatever came to their mind, and the researcher's intervention should be as minimal as possible and limited to neutral prompts when verbalization stopped for 5-7 seconds. Immediately after the TAP, each participant took part in the stimulated recall interview (SRI), focusing on selected moments from the reading session. The interview questions varied depending on the situations, such as why they took a long pause, why they skipped difficult words, or why they chose particular strategies. The TAP captured the participants' verbalized thoughts during reading, and SRI helped clarify moments in the TAP record that were difficult to interpret directly, such as long pauses, hesitation, skipped terms, or sudden changes in strategy (Ericsson & Simon, 1993; Pressley & Afflerbach, 1995; Gass & Mackey, 2000).

All recordings were transcribed verbatim, annotated with timestamps, and then segmented into the smallest analytic units, which represented a single idea or comprehension-related action. These segments were coded into three broad tiers: cue/obstacle, monitoring/appraisal, and regulatory strategy/response (see Table 1). Following that, related segments were grouped into comprehension episodes based on the sequence: cue → monitoring/appraisal → regulation → outcome, with outcomes categorized as recovered, partially recovered, or skipped/avoided. Pauses and code-switching were not treated automatically as evidence of difficulty, but interpreted in relation to surrounding contextual cues and, where necessary, the SRI explanation. In conceptual terms, this study treats reading comprehension as a process of meaning construction that can be disrupted and later restored through metacognitive monitoring and regulation (Flavell, 1979; Nelson & Narens, 1990).

© 2026 ACADEMY PUBLICATION

TABLE 1
OPERATIONAL CODEBOOK FOR CODING TAP

Tier	Code	Operational definition	Inclusion indicators	Exclusion/notes
Code1 (Cue/ Obstacle)	PAUSE_LONG	Silence ≥ 5–7 seconds plausibly linked to processing (tracked with phone timer).	[pause 7s], gaze fixed on text, pointing to line/term, silent rereading, effortful facial expression.	If a pause coincides with an external disruption (e.g., an entrance, loud noise), mark it as non-cognitive in the memo and do not code it as a processing cue.
	VOCA_BREAK	Lexical/terminological breakdown: difficulty understanding a word, term, abbreviation, or phrase.	"What does X mean?", "This term…?", Pronounce out a term, stopping at an abbreviation.	If difficulty is primarily about coherence/argument across sentences, use COH_BREAK.
	COH_BREAK	Coherence/argument breakdown: unclear relations across clauses/sentences or uncertainty about claims/evidence.	"This does not connect…," "Why is it like this?" "Where does this claim come from?"	It may occur without term difficulty; often followed by QUESTION or MONITOR_NEG.
	TRANSLATE/ CODESWITCH	Translation or language shifting (ID↔MK↔EN) is used in meaning-making. ID: Indonesian MK: Local dialect EN: English	Switch language to search for an equivalent, gloss, or continue reasoning.	Treat it as a candidate cue: interpret as a breakdown only if corroborated by other indicators and/or SRI confirms difficulty-triggered switching.
Code 2 (Monitoring/Appraisal)	MONITOR_NEG	Explicit negative appraisal of comprehension.	"I don't understand," "I am confused," "I'm not sure," "It doesn't make sense."	If the statement is evaluative but not about understanding (e.g., "this is important"), do not code it as MONITOR_NEG.
	QUESTION	An explicit question used to test/build understanding.	"What does it mean?" "Refers to what?" "So, is this saying…?"	Questions may signal monitoring even if no difficulty cue appears in the same segment.
	EVAL_CLAIM	Evaluating the credibility/logic of a claim or evidence.	"They say data show… but from where?", "This seems too general."	Use only if argumentative evaluation is part of the analytic focus.
Code 3 (Regulatory Strategy/ Response)	GLOSS_WORD	Creating an equivalent or short definition for a term/phrase.	"ICT means…," "LMS is like…," defining in simpler words.	Often follows VOCA_BREAK, but can occur proactively.
	PARAPHRASE	Restating the idea in one's own words to maintain coherence.	"So, the point is…," "It means that…"	If the participant repeats the sentence verbatim, do not code it as a paraphrase.
	INFER_CONNECT	Making an inference or connecting ideas across the text and/or prior knowledge.	"That implies…," "This links to the earlier point…," "So, in campus context…"	The participant must add a connection beyond surface restatement.
	REREAD	Rereading a portion of text (aloud or silently) to restore comprehension.	"Let me reread…," visible silent reread in video memo.	If rereading occurs due to a pronunciation slip, it is a non-strategy.
	SUMMARIZE	Summarizing a section to consolidate meaning before moving on.	"So, this paragraph is about…" short recap before continuing.	Distinct from paraphrase: summarize condenses multiple propositions.
	SKIP/AVOID	Skipping or avoiding an unclear part to proceed.	"I will skip this," "I will move on," continuing without resolution.	Often associated with outcome partial/skip; corroborate with SRI when possible.

III. FINDINGS AND DISCUSSION

A. Findings

The findings in this study are organized around comprehension episodes—the sequences of think-aloud segments that show how readers move from a difficulty cue to monitoring, then to regulatory action, and finally to an outcome. This approach allows us to go beyond listing isolated strategies to examine how comprehension was negotiated in real time. The original text and interview were in Bahasa Indonesia, but for this publication, we translated them into English.

© 2026 ACADEMY PUBLICATION

TABLE 2
COMPREHENSION EPISODES

Episode (Text Trigger)	Cue (Code1)	Monitoring (Code2)	Strategy (Code3)	Outcome	TAP Segment ID	Stimulated Recall (SR) Evidence
ICT / "information and communication technology (ICT)."	PAUSE_ LONG / TRANSLA TE	QUESTION / MONITOR_ NEG	GLOSS_ WORD + PARAPHRASE	Recovered	M2-018–M2-023	"The pause reflected an attempt to recall what ICT stands for before glossing it in Indonesian."
LMS / "Learning Management System (LMS)"	VOCA_ BREAK / PAUSE_L ONG	QUESTION	GLOSS_ WORD + PARAPHRASE (+ REREAD)	Recovered	M2-031–M2-036	"I glossed LMS so the paragraph would remain coherent and I could continue reading."
Disparity / "urban–rural disparity."	VOCA_ BREAK	QUESTION	GLOSS_ WORD + PARAPHRASE	Recovered	M1-053–M1-056	"I searched for an equivalent word—*kesenjangan*—to stabilize the meaning."
Digital accessibility	MONITOR _NEG / PAUSE_L ONG	QUESTION	GLOSS_WORD → INFER_CONNECT	Recovered	M1-044–M1-047	"I connected it to my own experience with signal coverage and data quotas."
Gamification / "points, badges, leaderboards."	VOCA_ BREAK	QUESTION	GLOSS_WORD + PARAPHRASE + INFER_CONNECT	Recovered	S3-060–S3-066	"I pictured it as a game-like reward system to make the concept concrete."
Evidence claim / "research indicates…"	(COH_BR EAK / MONITOR _NEG as cue)	QUESTION (source-oriented)	PARAPHRASE + SUMMARIZE (evaluative)	Recovered + evaluation	U3-040–U3-047	"I paused because the claim did not specify any evidence or citation."
Holistic / "a holistic approach."	VOCA_BR EAK	QUESTION	GLOSS_WORD + PARAPHRASE (+ INFER_CONNECT)	Recovered	M1-072–M1-076	"Later I realized the phrase framed the section as a set of integrated solutions."

Note: Five digits in segment ID consists of the participant's code (e.g., M1) and the order of segmentation (e.g., 036)

Across the dataset, episodes with the most evidence were concentrated around three types of textual triggers: technical or foreign terms and abbreviations (e.g., *ICT*, *LMS*, *VR*), abstract concepts (e.g., *disparity*, *digital accessibility*, *holistic*), and evidential claims (e.g., *data show…*, *research indicates…*). Analysis of these episodes revealed three recurring patterns: repair chains, concretization bridges, and differentiated monitoring. In Table 2, stimulated recall data were used selectively to clarify the meaning of pauses, language shifts, and strategy choices in key episodes. The following sections elaborate on the three patterns.

1. Repair Chains as a Recurrent Mechanism of Recovery

The repair chain refers to the relatively stable sequence in which comprehension breakdown was followed by monitoring and one or more regulatory moves that restored meaning. In these episodes, recovery was not contingent upon a single strategy but through a series of small but connected actions, such as pausing, questioning, glossing, paraphrasing, rereading, or linking the unfamiliar item to the immediate context.

This pattern was especially visible in episodes triggered by unfamiliar or technical terms. For instance, when participants encountered items such as *ICT*, *LMS*, or *disparity*, the episode often began with a cue such as hesitation, a prolonged pause, lexical breakdown, or language shifting. This condition was then followed by explicit monitoring, typically in the form of a question or a negative appraisal of understanding. Regulatory action usually took the form of generating a simpler equivalent, paraphrasing the surrounding clause, or briefly rereading the sentence before continuing.

a) Data 1
Episode "ICT" (S1)

The episodes began as the participant read a phrase about ICT
S1-024: *"ICT...*" coded as VOCA_BREAK

S1-025: "[silent for 6 seconds]" ---coded as PAUSE_LONG

S1-026: *"ICT... I don't know"* coded as VOCA_BREAK+MONITOR_NEG
In this episode, S1 met unfamiliar terms then followed by appraisal move but still didn't know and chose to skip. In SRI S1 stated that the tendency to skip because he was uncertain what ICT stands for.

b) Data 2
Episode "Disparity" (M1)

© 2026 ACADEMY PUBLICATION

In the sentence "data show a large disparity in internet access between urban and rural areas,"
The episode began with a difficulty cue in form of lexical breakdown:
M1-053: *"disparity…eee"* coded as VOCA_BREAK

An appraisal immediately followed this cue in the form of an explicit question marking a comprehension gap.
M1-054: *"What does "disparitas" mean?"* coded as QUESTION

After that, a meaningful pause occurred.
M1-055: —"[silent for 6 seconds]" ---coded as PAUSE_LONG
It signaled a processing space prior to the regulatory action.

M1-056: *"Disparity means a gap/large difference"* coded as REPAIR_STRAT + GLOSS_WORD + PARAPHRASE

After identifying an appropriate equivalent, the participant resumed reading more steadily and, in subsequent segments, was able to elaborate consequences (e.g., linking slow connectivity to disrupted learning). Stimulated recall supported these interpretations by showing that such pauses and shifts often occurred as a processing space for lexical retrieval or meaning consolidation rather than as signs of distraction.

Varied density and efficient use of repair chains were apparent based on students' levels of Indonesian proficiency (UKBI). Semenjana (basic level) participants tended to produce longer, less stable chains, often with repeated questions and more fillers, and ended with partial or no outcomes. Madya (intermediate level) participants showed a higher frequency of complete and reasonably efficient repair chains, combined with glossing and paraphrasing to restore local coherence. Unggul (proficient) participants generally displayed more compact and sometimes preventive chains, stabilizing meaning quickly before breakdowns became prolonged. In some cases, the repair went beyond word meaning to the level of argument and evidence.

2. Concretization Bridges in Recovering Abstract Meaning

The concretization bridge allows readers to recover abstract concepts by connecting them to examples, familiar situations, or lived experience. In this study, the concretization bridge was most evident when participants encountered terms that require more than a denotation, such as "digital accessibility", "inclusive learning", or "gamification". In these cases, readers often moved from uncertainty to understanding by anchoring the concept in a concrete situation they could imagine.

For example, when discussing "digital accessibility", some participants first expressed uncertainty through hesitation or direct questioning, then attempted a provisional gloss, and linked the term to practical situations such as poor signal coverage, a lack of devices, or difficulties with data quotas.

a) Data 3
Episode "Digital Accessibility" (M1)

The episode began as the participant read a phrase about limited connectivity *and barriers to digital accessibility*.
M1-043: "*barriers to digital accessibility*" was coded as reading segment.
When the abstract term appeared, the participant marked a processing disruption through negative monitoring:

M1-044: *"digital accessibility… hmm…"* coded as MONITOR_NEG
which then developed into an explicit question

M1-045: *"What does digital accessibility mean?"* coded as QUESTION
Following this appraisal, the participant engaged in regulation by producing an operational definition

M1-046: *"Maybe it is about whether you can access the internet/devices"* coded as REPAIR_STRAT + GLOSS_WORD + PARAPHRASE
Notably, the recovery did not stop at the definition. M1 immediately linked the concept to a concrete consequence grounded in a learning experience.

M1-047: *"This makes the quality of learning people receive unequal"* coded as REPAIR_STRAT + PARAPHRASE + INFER_CONNECT

Here, inference functioned as a "bridge" that tied the abstract concept (accessibility) to an easily imagined situation model (unequal learning quality), enabling the participant to continue reading with improved coherence. Stimulated recall strengthened this interpretation by showing that participants themselves experienced such examples as a way of "making the term real" so that the next part of the sentence would make sense.

b) Data 4

© 2026 ACADEMY PUBLICATION

Episode “Interactive Simulation” (S2)

The episodes began when participant read about the use of simulation application.
S2-088: *“Interactive simulation.. I’m not really understand”* coded as VOCA_BREAK+MONITOR_NEG

S2-089: *“Wait, maybe this means practicing independently using a computer”* coded as REPAIR_STRAT + PARAPHRASE + INFER_CONNECT.

Here the participant try to recover through highly concrete references as a way to grasp the concept.

For Semenjana (basic level) participants, concretization often served as the primary route to understanding, but it sometimes yielded only situational familiarity rather than a fully stable conceptual definition. Madya participants often combined concretization with initial glossing or paraphrasing to produce a more secure recovery. Among Unggul (proficient) participants, concretization tended to make the concept imaginable while supporting conceptual expansion and implication-making. In these cases, examples were used not only to understand the concept but also to test its broader relevance and consequences within the text argument.

3. From Meaning Monitoring to Epistemic Monitoring

The third and particularly important pattern was the nature of monitoring itself. Our data showed that monitoring not only acknowledged, “I do not understand,” but also appeared in at least two forms. The first was meaning monitoring, in which participants noticed that a word, term, or relationship between ideas was unclear. In the second form, epistemic monitoring, participants questioned whether a claim was sufficiently supported, the evidence was specified, or the argument's structure was convincing.

In this study, meaning monitoring appeared across all levels, especially around unfamiliar terms and abstract concepts. Participants often expressed meaning through direct questions, hesitation, or negative appraisal, and this was usually followed by glossing, paraphrasing, or inference. Below are the examples:

a) Data 5
Episode “transformation” (M1)

In the portion of text stating about the transformation of learning practices, M1 showed instability:
M1-008: *“transformation… hmm…”*coded as *MONITOR*_NEG

This monitoring was then intensified through an explicit question:
M1-009: *“What does transformation mean here?”* coded as QUESTION

M1-010: —“[silent for 6 seconds]” ---coded as PAUSE_LONG

After a pause, the participant enacted regulation through paraphrase and inference:
M1-011: *"Maybe it means a change in the way people learn, becoming more digital, maybe”* coded as REPAIR_STRAT+PARAPHRASE+ INFER_CONNECT

b) Data 6
Episode “Data Show..” (U3)

When participant met the section like “Data show”
U3-041: *“Here is stated that Data show… but there is no sources mention..hmm”* coded as MONITOR_NEG
U3-042: *“The sources must be stated, maybe BPS or Kemenkoinfo?)* coded as REPAIR_STRAT + INFER_CONNECT

c) Data 7
Episode “Research indicates”

The episodes began when participant read phrase like “research indicates”
U3-056: *“Research indicates..this is a general claim, so it should include a reference to make it stronger”* coded as MONITOR_NEG.

U3-057: *“Which research?”* coded as QUESTION

U3-058: *“But the concept is gamification increases engagement through rewards”* coded as REPAIR_STRAT+PARAPHRASE+INFER_CONNECT
In SR, the participant says that he paused because the claim did not specify any evidence or citation.

© 2026 ACADEMY PUBLICATION

The sequence above illustrates monitoring as an alarm that signals a meaning gap and simultaneously guides the form of repair, which, in this case, is the concept-based paraphrasing.

Epistemic monitoring, however, was most visible among participants in *the Unggul* level (proficient). Some of them understood episodes containing phrases like 'research indicates' or 'data show…' but questioned whether the claim had enough evidence and a strong connection with supporting evidence, or lacked citations. The subsequent stimulated recall confirmed that these responses reflected an evaluative stance rather than simple lexical or conceptual confusion.

Across the case, monitoring emerged as a central node linking obstacles to regulation, but the object of monitoring differed by proficiency level. *Semenjana* participants (intermediate level) showed intensive meaning monitoring, although it did not consistently lead to a stable recovery, whereas *Madya* participants tended to transform monitoring more directly into productive repair. *Unggul* participants extended monitoring into the epistemic domain, showing that academic reading in this dataset involved not only constructing meaning but also evaluating the warrant of what was being read.

B. Discussion

The findings of this research indicate that academic reading is best understood not as a smooth, linear act of processing, but as a process negotiated through recurring episodes of disruption and recovery. This study has demonstrated that the participants build comprehension through local moments when they meet a difficult part, assess it, and respond through regulation. These findings are consistent with the view that reading comprehension is the construction and maintenance of a coherent mental representation of text, which must be actively sustained during lexical, conceptual, or argumentative difficulties (Kintsch, 1998). In this sense, difficulty is not peripheral to academic reading; it is part of the process through which understanding is formed. It aligns with more recent work arguing that reading comprehension should be understood as a dynamic process shaped by interactions between reader characteristics, text features, and context rather than the product of isolated subskills alone (Duke & Cartwright, 2021; McMaster & Kendeou, 2023).

The first finding about the repair chain supports and extends metacognitive accounts of reading. The sequence observed in the data closely resembles the distinction between monitoring and control proposed in metacognitive theory (Flavell, 1979; Nelson & Narens, 1990). Participants first registered a disruption, then appraised the problem, and finally selected one or more actions to restore coherence. It is important to note that regulation rarely appeared as a single, isolated strategy, but through layers of actions that involve pausing, questioning, glossing, paraphrasing, rereading, and inferencing. This interpretation is aligned with a previous study by Tibken, who reported that comprehension monitoring and regulation among university students in Germany are especially important when they engage with expository texts and conceptually difficult material (Tibken & Tiffin-Richards, 2025). In this sense, the study extends current understandings of reading strategies by showing that what matters is not only which strategies are used, but how they are organized into a recoverable mechanism that can be empirically traced.

The second finding, the concretization bridge, shows that abstract concepts were often understood not from formal definition alone, but through movement toward concrete experience or familiar scenarios. This finding supports previous research on self-explanation, which reported that readers deepen their understanding when restating ideas in meaningful, personally interpretable ways (McNamara, 2004) and that verbalized responses during reading can support coherence building and reveal differences in how readers construct mental representations of the text (McCrudden et al., 2024). This resonates with the constructionist's ideas of comprehension that readers generate inferences to maintain coherence and build a workable mental model of the text (Graesser et al., 1994). In this study, however, the concretization bridge is not simply explanatory but specifically vital to stabilize meaning when facing difficulties. The findings also reveal an important distinction across proficiency levels. Concretization practiced by readers with moderate and higher proficiency can support recovery, whereas those at the lower proficiency level may not consistently produce a stable conceptual understanding. In this sense, a participant can get the example but not necessarily grasp the underlying abstraction.

The third finding, differentiated monitoring, is arguably the strongest conceptual contribution of this study. By distinguishing between meaning monitoring and epistemic monitoring, this study suggests that academic reading involves understanding of what a text says, as well as evaluating how well the claims in the text are supported. Accordingly, this finding expands the comprehension of academic reading beyond vocabulary and local coherence levels to knowledge appraisal, especially when readers encounter generalized or weak claims. In this respect, the findings resonate with the work on epistemic cognition and document processing, where beliefs about justification shape how readers evaluate and use information (Bråten et al., 2022). Although this study involved a single academic text rather than multiple documents, the same orientation was visible. Some participants were disrupted not because they could not decode the sentence, but because they questioned its evidential adequacy. Recent work on comprehension monitoring also supports the idea that the text characteristics can influence the monitoring accuracy; therefore, readers should check their monitoring process repeatedly (Hildenbrand et al., 2024; Tibken & Tiffin-Richards, 2025).

The variation across Indonesian proficiency levels (UKBI) reinforces this interpretation. Differences are not primarily about whether participants are monitored or regulated, but rather what they monitor, how efficiently they exercise regulation, and how far the outcome extends. At the *Semenjana* (basic level), monitoring was often intense, but regulation remained fragile, sometimes ending in partial recovery or avoidance, while students in *Madya* (intermediate level) readily converted monitoring into effective repair. The *Unggul* (proficient) students used regulation more economically and monitoring more evaluatively, extending to the level of argument quality and justification. These patterns are consistent with the process-oriented reading research, suggesting that comprehension differences are reflected in the final

© 2026 ACADEMY PUBLICATION

performance as well as the quality of online processing, reanalysis, and coherence-building activity during reading (McCrudden et al., 2024; Tibken & Tiffin-Richards, 2025).

Taken together, the three findings suggest that academic reading at the university level involves the interplay of cognitive, metacognitive, and evaluative work. When students are expected to engage with and critically evaluate complex texts containing many abstract concepts and claims, reading cannot be understood merely as the extraction of meaning from text. The repair chain explains how students procedurally restore comprehension when meaning breaks down, the concretization bridge shows how abstract ideas become workable through experiential anchoring, and the differentiated monitoring demonstrates that academic reading includes not only sense-making but also claim evaluation. These mechanisms support an episode-based account of academic reading in which meaning is not simply extracted from the text, but actively built, stabilized, and, at times, questioned. This study highlights how university students manage the cognitive and evaluative demands of academic reading. It offers a broader view that is compatible with recent calls to align theories of reading more closely with the dynamic ways in which readers actually interact with texts in context (McMaster & Kendeou, 2023).

IV. Conclusion

This study has shown that academic reading is more accurately understood as a series of comprehension episodes through which readers negotiate difficulty, monitor understanding, regulate meaning, and, at times, evaluate the adequacy of textual claims. Across the cases, three salient mechanisms were identified: repair chains to restore comprehension, concretization bridges to help readers stabilize abstract meaning through concrete examples and experience, and differentiated monitoring to conduct meaning-focused appraisal and epistemic evaluation of evidence and argument. These findings suggest that academic reading in higher education involves more than decoding and meaning construction, but requires the students to manage uncertainty, sustain coherence, and judge whether the text contains enough support for the claims.

Conceptually, the study contributes an episode-based account of academic reading that integrates cognitive, metacognitive, and evaluative processes within a single analytic framework. Methodologically, it demonstrates the value of combining the think-aloud protocol (TAP) with the stimulated recall interview (SRI) to build a more transparent chain of evidence for interpreting reading processes. As a qualitative case study, this research presents an analytically grounded account of how academic reading unfolds in a linguistically diverse and underrepresented higher-education context. The limitation of this study is the lack of statistical generalization and the use of small samples. Future research can involve participants from a wider range of universities and disciplines to examine whether the identified episode patterns remain consistent across contexts. The use of different types of academic texts may lend insights into how textual demands shape comprehension breakdown, monitoring, and regulation.

Acknowledgements

The researchers would like to thank Lembaga Pengelola Dana Pendidikan (LPDP/ The Indonesia Endowment Fund for Education), Indonesian Education Scholarship (BPI), and PPAPT (Pusat Pembiayaan dan Asesmen Pendidikan Tinggi) for sponsoring the publication of this research.

References

[1] Abas, I. H., & Abd Aziz, N. H. (2018). Model of the writing process and strategies of EFL proficient student writers: A case study of Indonesian learners. *Pertanika Journal of Social Sciences & Humanities*, *26*(3), 1815–1842.

[2] Adams, M. J. (1990). *Beginning to Read: Thinking and Learning about Print*. MIT Press.

[3] Alruthaya, S., Mantei, J., White, S. L. J., & Kervin, L. (2025). Moving towards a more comprehensive understanding of multimodal text complexity. *The Australian Journal of Language and Literacy*, *48*(2), 119–139. https://doi.org/10.1007/s44020-025-00079-9

[4] Atkinson, D., & Smith Risser, H. (2023). Exploring students' responses to reading assignments in first-year university mathematics courses. *Journal of College Reading and Learning*, *53*(1), 38–57. https://doi.org/10.1080/10790195.2022.2126415

[5] Bergman, L. (2024). Students' reading in higher education: Challenges and ways forward. *Journal of Adolescent & Adult Literacy*, *67*(6), 414–423. https://doi.org/10.1002/jaal.1346

[6] Bowles, M. A. (2018). Introspective verbal reports: Think-alouds and stimulated recall. In A. Phakiti, P. De Costa, L. Plonsky, & S. Starfield (Eds.), *The Palgrave handbook of applied linguistics research methodology* (pp. 339–357). Palgrave Macmillan. https://doi.org/10.1057/978-1-137-59900-1_16

[7] Bråten, I., Brandmo, C., Ferguson, L. E., & Strømsø, H. I. (2022). Epistemic justification in multiple document literacy: A refutation text intervention. *Contemporary Educational Psychology*, *71*, Article 102122. https://doi.org/10.1016/j.cedpsych.2022.102122

[8] Chall, J. S. (1983). *Stages of Reading Development*. McGraw-Hill.

[9] Chapman, M., Quick, S. N., & Freeman, P. (2024). An investigation of in situ cognitions of English Premier League academy football coaches using stimulated recall and think aloud protocol. *Journal of Coaching and Sports Science*, *3*(2), 99–112. https://doi.org/10.58524/002024335600

[10] Charters, E. (2003). The use of think-aloud methods in qualitative research: An introduction to think-aloud methods. *Brock Education Journal*, *12*(2), 68–82. https://doi.org/10.26522/brocked.v12i2.38

© 2026 ACADEMY PUBLICATION

[11] Craig, K., Hale, D., Grainger, C., & Stewart, M. E. (2020). Evaluating metacognitive self-reports: Systematic reviews of the value of self-report in metacognitive research. *Metacognition and Learning*, *15*, 155–213. https://doi.org/10.1007/s11409-020-09222-y
[12] Duke, N. K., & Cartwright, K. B. (2021). The science of reading progresses: Communicating advances beyond the simple view of reading. *Reading Research Quarterly*, *56*(S1), S25–S44. https://doi.org/10.1002/rrq.411
[13] Ericsson, K. A., & Simon, H. A. (1993). *Protocol Analysis: Verbal Reports as Data*. MIT Press.
[14] Fan, Y., Rakovic, M., van der Graaf, J., Lim, L., Singh, S., Moore, J., Molenaar, I., Bannert, M., & Gašević, D. (2023). Towards a fuller picture: Triangulation and integration of the measurement of self-regulated learning based on trace and think-aloud data. *Journal of Computer Assisted Learning*, *39*(4), 1303–1324. https://doi.org/10.1111/jcal.12801
[15] Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*, *34*(10), 906–911.
[16] Gass, S. M., & Mackey, A. (2000). *Stimulated recall methodology in second language research*. Lawrence Erlbaum Associates.
[17] Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, *101*(3), 371–395. https://doi.org/10.1037/0033-295X.101.3.371
[18] Hildenbrand, L., Sarmento, D., Griffin, T. D., & Wiley, J. (2024). Conceptual overlap among texts impedes comprehension monitoring. *Psychonomic Bulletin & Review*, *31*(2), 750–760. https://doi.org/10.3758/s13423-023-02349-4
[19] Hoang, G. T. L., & Storch, N. (2024). How do EFL learners process and uptake Criterion automated corrective feedback? Insights from two case studies. *Iranian Journal of Language Teaching Research*, *12*(1), 1–20. https://doi.org/10.30466/ijltr.2024.121420
[20] Holschuh, J. P. (2020). College Reading and studying: The complexity of academic literacy task demands. *Journal of Adolescent & Adult Literacy*, *62*(6), 599–604. https://doi.org/10.1002/jaal.876
[21] Hung, C. O.-Y., Cai, M., & Liao, X. (2025). Contribution of executive function to different levels of reading comprehension. *Journal of Research in Reading*. https://doi.org/10.1111/1467-9817.70011
[22] Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge University Press.
[23] Kiili, C., Strømsø, H. I., Bråten, I., Ruotsalainen, J., & Räikkönen, E. (2024). Reading comprehension skills and prior topic knowledge serve as resources when adolescents justify the credibility of multiple online texts. *Reading Psychology*, *45*(7), 662–689. https://doi.org/10.1080/02702711.2024.2351485
[24] Lea, M., & Street, B. (1998). Student writing in higher education: An academic literacies approach. *Studies in Higher Education*, *23*(2), 157–172. https://doi.org/10.1080/03075079812331380364
[25] Mason, W., & Warmington, M. (2024). Academic reading as a grudging act: How do higher education students experience academic reading and what can educators do about it? *Higher Education*, *88*, 839–856. https://doi.org/10.1007/s10734-023-01145-2
[26] McCrudden, M. T., Huynh, L., Lyu, B., Kulikowich, J. M., & McNamara, D. S. (2024). Coherence building while reading multiple complementary documents. *Contemporary Educational Psychology*, *77*, Article 102266. https://doi.org/10.1016/j.cedpsych.2024.102266
[27] McMaster, K. L., & Kendeou, P. (2023). Refocusing reading comprehension: Aligning theory with assessment and intervention. *Learning and Individual Differences*, *102*, Article 102256. https://doi.org/10.1016/j.lindif.2023.102256
[28] McNamara, D. S. (2004). SERT: Self-explanation reading training. *Discourse Processes*, *38*(1), 1–30. https://doi.org/10.1207/s15326950dp3801_1
[29] Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 26, pp. 125–173). Academic Press. https://doi.org/10.1016/S0079-7421(08)60053-5
[30] Organization for Economic Co-operation and Development. (2023). *PISA 2022 results (Volume I and II): Country note—Indonesia*. OECD.
[31] Paige, D., Rupley, W. H., & Ziglari, L. (2024). Critical thinking in reading comprehension: Fine-tuning the simple view of reading. *Education Sciences*, *14*(3), Article 225. https://doi.org/10.3390/educsci14030225
[32] Paulhus, D. L. (2017). Socially desirable responding on self-reports. In V. Zeigler-Hill & T. Shackelford (Eds.), *Encyclopedia of personality and individual differences*. Springer. https://doi.org/10.1007/978-3-319-28099-8_1349-1
[33] Perfetti, C. A. (1985). *Reading ability*. Oxford University Press.
[34] Pressley, M., & Afflerbach, P. (1995). *Verbal protocols of reading: The nature of constructively responsive reading*. Lawrence Erlbaum Associates.
[35] Rayner, K., Pollatsek, A., Ashby, J., & Clifton, C. (2012). *The Psychology of Reading* (2nd ed.). Routledge.
[36] Reynolds, T., Rush, L. S., Lampi, J. P., & Holschuh, J. P. (2020). English disciplinary literacy: Enhancing students' literary interpretive moves. *Journal of Adolescent & Adult Literacy*, *64*(2), 201–209. https://doi.org/10.1002/jaal.1066
[37] Schellings, G., & Van Hout-Wolters, B. H. A. M. (2011). Measuring strategy use with self-report instruments: Theoretical and empirical considerations. *Metacognition and Learning*, *6*(2), 83–90. https://doi.org/10.1007/s11409-011-9081-9
[38] Smith, F. (2004). *Understanding reading: A psycholinguistic analysis of reading and learning to read*. Routledge.
[39] Snow, C. E., & Sweet, A. P. (2003). Reading for comprehension. In A. Sweet & C. Snow (Eds.), *Rethinking reading comprehension* (pp. 1–11). Guilford Press.
[40] Supardan, A. I., & Capati, M. W. (2026). AI as a metacognitive mirror: How students use AI to monitor and repair reading comprehension breakdowns. *DUTIES: Education and Humanities International Journal*, *2*(1), 1–24. https://doi.org/10.70152/duties.v2i1.320
[41] Tang, K., Noordin, N., & Ismail, L. (2025). Evaluating the think-aloud method for English reading: A systematic review. *Cogent Education*, *12*(1), Article 2497150. https://doi.org/10.1080/2331186X.2025.2497150
[42] Tian, L., Liu, Q., & Zhang, X. (2022). Self-regulated writing strategy use when revising upon automated, peer, and teacher feedback in an online English as a foreign language writing course. *Frontiers in Psychology*, *13*, Article 873170. https://doi.org/10.3389/fpsyg.2022.873170

© 2026 ACADEMY PUBLICATION

[43] Tibken, C., & Tiffin-Richards, S. P. (2025). Reading behavior as an indicator of comprehension monitoring when reading expository texts. *Metacognition and Learning*, *20*, Article 38. https://doi.org/10.1007/s11409-025-09440-2
[44] Veenman, M. V. J. (2011). Alternative assessment of strategy use with self-report instruments: A discussion. *Metacognition and Learning*, *6*(2), 205–211. https://doi.org/10.1007/s11409-011-9080-x
[45] Zhai, X., Chu, X., Wang, M., Tsai, C.-C., & Spector, J. M. (2024). A systematic review of stimulated recall (SR) in educational research from 2012 to 2022. *Humanities and Social Sciences Communications*, *11*, Article 489. https://doi.org/10.1057/s41599-024-02987-6

Lenny Nofriyani Adam is a Ph.D. candidate at Universitas Pendidikan Indonesia, Bandung, Indonesia. She is a lecturer at Department of Indonesian Language and Literature Education, Faculty of Education, Universitas Timor, Indonesia. Her research is related to Language Acquisition, Language Comprehension, Psycholinguistics, and Pragmatics issues. ORCID iD: https://orcid.org/0000-0002-4715-1426

Dadang Sudana is a lecturer at the Linguistics Study Program, School of Post Graduate, Universitas Pendidikan Indonesia, Bandung, Indonesia. His research interests are focused on Language acquisition, Morphology and Semantics. ORCID iD: https://orcid.org/0000-0001-5370-1685

Heru Pratikno is a Ph.D. candidate at Universitas Pendidikan Indonesia, Bandung, Indonesia. He is a lecturer at Universitas Islam Bandung, Indonesia. His research is related to Morphology, Language Politeness, Humor, and Pragmatics issues. ORCID iD: https://orcid.org/0000-0001-7704-208x

Joni Soleman Nalenan is a Ph.D. candidate at Universitas Pendidikan Indonesia, Bandung, Indonesia. He is a lecturer at Department of Indonesian Language and Literature Education, Faculty of Education, Universitas Timor, Indonesia. His research is related to Critical Discourse Analysis, Cultural Linguistics, Anthropolinguistics and Syntax. ORCID iD: https://orcid.org/0000-0001-8822-0101

Yandris Obligus Nabuasa is a Master candidate at Universitas Pendidikan Indonesia, Bandung, Indonesia. He is an English Teacher at Banat Senior High School (SMA Negeri Banat), East Nusa Tenggara, Indonesia. His research is related to English Teaching, General and Character Education and culture of South-Central Timor. ORCID iD: https://orcid.org/0009-0007-6934-0969

Hanny Withney Gloria Wayangkau is a Master candidate at Universitas Pendidikan Indonesia, Bandung, Indonesia. She is an English teacher at Sekolah Negeri Khusus Olahraga (Sport Senior High School), Jayapura – Papua, Indonesia. Her research is related to Curriculum and Materials Development, Language and Literature Education, and Linguistic Development. ORCID iD: https://orcid.org/0009-0006-1586-7282

Fauzan Novaldy Pratama is a Ph.D. candidate at Universitas Pendidikan Indonesia, Bandung, Indonesia. He is a lecturer at Department of English Language (Vocational/D3), Faculty of Literature, Pasim National University, Indonesia. His research is related to Pragma-Semantics, Computational Linguistics, and Critical Discourse Analysis. ORCID iD: https://orcid.org/0009-0002-6312-7597

© 2026 ACADEMY PUBLICATION