

## Epistemic Agency in Multimodal Reading: How Primary Learners Construct Knowledge Across Modes

**Hasna Fitri Labibah\***

\*Doctoral Program of Language Education, Faculty of Languages and Arts,  
Universitas Negeri Semarang, Indonesia  
E-mail: canachan77@students.unnes.ac.id

**Januarius Mujiyanto\*\***

\*\*Doctoral Program of Language Education, Faculty of Languages and Arts,  
Universitas Negeri Semarang, Indonesia  
E-mail: yanmujiyanto@mail.unnes.ac.id

**Dwi Rukmini\*\*\***

\*\*\*Doctoral Program of Language Education, Faculty of Languages and Arts,  
Universitas Negeri Semarang, Indonesia  
E-mail: wiwidwirukmini71@gmail.com

**Widhiyanto\*\*\*\***

\*\*\*\*Doctoral Program of Language Education, Faculty of Languages and Arts,  
Universitas Negeri Semarang, Indonesia  
E-mail: pakwidhi\_english@mail.unnes.ac.id

Received: August 19<sup>th</sup>, 2025. Accepted: October 10<sup>th</sup>, 2025. Published: October 31<sup>st</sup>, 2025.

### Abstract

Reading literacy remains a persistent challenge in many English as a Foreign Language (EFL) contexts. However, primary EFL learners frequently encounter difficulties in integrating linguistic, visual, spatial, and gestural modalities. This study investigates how primary learners exercise epistemic agency in multimodal reading and construct knowledge across different modes. Utilizing multimodality theory and epistemic agency, the study implemented an explanatory sequential mixed-methods design involving thirty fifth-grade students from Madrasah Ibtidaiyah Negeri (MIN) 1 Tulang Bawang Barat, Lampung. In Phase 1, a one-group pretest–posttest design measured changes in multimodal comprehension using a validated twenty-item instrument. Results showed significant improvement, with mean scores rising from 61.40 (SD = 8.25) to 74.30 (SD = 7.80),  $t(29) = 8.52$ ,  $p < 0.001$ , and Cohen's  $d = 1.56$ , indicating a large effect. In Phase 2, six students representing varied improvement levels were interviewed and engaged in task-based reflections. Thematic analysis identified five dimensions of epistemic agency: taking initiative, decision-making across modes, justification of understanding, reflection and self-regulation, and collaboration and dialogue. The integration of both phases demonstrated that learners' agentive engagement with multimodal resources underpinned comprehension gains, transforming affordances into strategies for meaning-making. These findings affirm that multimodal pedagogy fosters not only measurable comprehension outcomes but also autonomy, critical thinking, and collaborative engagement. The study recommends embedding multimodal literacy

into EFL curricula and providing scaffolds that balance cognitive support with opportunities for agency.

**Keywords:** *epistemic agency, knowledge construction, literacy pedagogy, multimodal reading, primary school.*

### Abstrak

Literasi membaca tetap menjadi tantangan yang terus-menerus dalam banyak konteks Bahasa Inggris sebagai Bahasa Asing (EFL). Namun, pembelajaran EFL tingkat dasar sering menghadapi kesulitan dalam mengintegrasikan modalitas linguistik, visual, spasial, dan gestural. Studi ini menyelidiki bagaimana pembelajaran primer menjalankan agensi epistemik dalam membaca multimodal dan membangun pengetahuan di berbagai moda. Dengan memanfaatkan teori multimodalitas dan agensi epistemik, penelitian ini menerapkan desain campuran sekuensial eksplanatif yang melibatkan tiga puluh siswa kelas lima dari Madrasah Ibtidaiyah Negeri (MIN) 1 Tulang Bawang Barat, Lampung. Pada Fase 1, desain pretes-postes satu kelompok mengukur perubahan dalam pemahaman multimodal menggunakan instrumen dua puluh item yang tervalidasi. Hasil menunjukkan peningkatan yang signifikan, dengan skor rata-rata meningkat dari 61,40 (SD = 8,25) menjadi 74,30 (SD = 7,80),  $t(29) = 8,52$ ,  $p < 0,001$ , dan Cohen's  $d = 1,56$ , yang menunjukkan efek yang besar. Pada Fase 2, enam siswa yang mewakili berbagai tingkat peningkatan diwawancara dan dilibatkan dalam refleksi berbasis tugas. Analisis tematik mengidentifikasi lima dimensi agensi epistemik: mengambil inisiatif, pengambilan keputusan lintas moda, pemberanakan pemahaman, refleksi dan pengaturan diri, serta kolaborasi dan dialog. Integrasi kedua fase menunjukkan bahwa keterlibatan agen pembelajaran dengan sumber daya multimoda mendukung perolehan pemahaman, mengubah affordance menjadi strategi untuk pembuatan makna. Temuan ini menegaskan bahwa pedagogi multimoda tidak hanya mendorong hasil pemahaman yang terukur tetapi juga otonomi, pemikiran kritis, dan keterlibatan kolaboratif. Studi ini merekomendasikan penanaman literasi multimodal ke dalam kurikulum EFL dan menyediakan peran yang menyeimbangkan dukungan kognitif dengan peluang untuk agensi.

**Kata kunci:** *agensi epistemik, konstruksi pengetahuan, pedagogi literasi, membaca multimodal, sekolah dasar.*

## INTRODUCTION

Reading literacy has increasingly been recognized as a multidimensional practice that extends beyond the decoding of printed text to the orchestration of diverse semiotic resources. Research on multiliteracies underscores that learner must navigate multimodal environments where meaning is conveyed through linguistic, visual, spatial, gestural, and digital resources (Cope & Kalantzis, 2020; Fälth et al., 2023; Jewitt, 2014). In EFL contexts, this challenge is particularly urgent, as students are expected not only to acquire linguistic competence but also to mobilize multimodal cues to support comprehension (Rohi & Nurhayati, 2024; Yi, Dong, et al., 2024; Yi, Zhao, et al., 2024). International assessments such as PISA 2022, have shown persistent gaps in reading performance, especially in developing countries, highlighting the need for more inclusive and innovative literacy pedagogies (AlAli & Wardat, 2024; Mahapoonyanont & Songsang, 2024; Norman, 2023).

In response, scholars have turned their attention to multimodal literacy as a framework for rethinking reading pedagogy. Multimodality views meaning-making as distributed across linguistic, visual, audio, gestural, and spatial resources rather than confined to print alone (Jewitt, 2014; Mills, 2016). Such perspectives challenge traditional text-based instruction by emphasizing the orchestration of multiple semiotic modes to enhance comprehension and engagement. Studies report that multimodal designs engage learners more deeply, foster critical and inferential comprehension, and bridge gaps for those who struggle with linguistic texts (January & Nima, 2018; Weninger, 2023). Evidence suggests that multimodal approaches also promote equity by allowing diverse learners to access meaning through multiple entry points (Daulay & Dewi, 2025; Santoso & Mangkuluhur, 2024; White, 2024). However, multimodality is not only about exposure to multiple resources but also about how learners actively use them. Scholars argue that literacy pedagogy must foreground agency, positioning learners as decision-makers who strategically select, evaluate, and justify their use of multimodal resources (Biesta et al., 2015; Van de Pol et al., 2019).

This emphasis on epistemic agency represents a critical shift in literacy research. Rather than viewing learners as passive consumers of multimodal texts, recent work conceptualizes them as epistemic agents who take responsibility for advancing their own understanding and contributing to collective meaning-making (Rappa & Tang, 2018; Yang, 2019). Empirical studies highlight how agency emerges in tasks that require learners to reflect, collaborate, and regulate their strategies (Biesta et al., 2015; Van de Pol et al., 2019). In multimodal environments, agency becomes especially salient, as learners must make choices across modes, negotiate tensions, and justify their interpretations (Vaughn et al., 2020; Zhang et al., 2022).

Despite this growing body of literature, several gaps remain. Much existing research on multimodal literacy has been conducted in secondary or tertiary contexts, often focusing on digital or academic literacy practices (Mills, 2016; White, 2024). Studies with younger learners, particularly in primary EFL settings, remain limited, even though early interventions are crucial for developing lifelong literacy habits (Daulay & Dewi, 2025; Santoso & Mangkuluhur, 2024). Furthermore, while multimodal approaches have been shown to improve comprehension, fewer studies explicitly investigate the role of epistemic agency in mediating these gains (Vaughn et al., 2020; Zhang et al., 2022). There is still insufficient evidence on how learners in primary schools navigate multimodal resources and how agency manifests in their comprehension processes.

Another gap lies in methodological approaches. Previous studies often rely on either quantitative assessment of reading gains or qualitative explorations of classroom practices, but rarely integrate both systematically (Kim et al., 2025; Vaughn et al., 2020; Yang, 2019). As a result, the field lacks comprehensive evidence that not only demonstrates measurable improvements in comprehension but also explains the agentive processes underpinning them. Mixed-methods designs, particularly explanatory sequential approaches, are well-suited to address this gap by linking quantifiable outcomes with in-depth qualitative accounts (Creswell & Plano Clark, 2023; Zhang et al., 2022).

Addressing these gaps, the present study investigates how primary EFL learners exercise epistemic agency in multimodal reading and how they construct knowledge across modes. Specifically, it examines both the measurable impact of multimodal reading on

comprehension and the ways in which learners demonstrate initiative, decision-making, justification, reflection, and collaboration. By situating the study in an Indonesian primary school context, it responds to calls for more research in underrepresented settings where literacy achievement remains a national concern (Daulay & Dewi, 2025; Januarty & Nima, 2018; Santoso & Mangkuluhur, 2024).

The study contributes to the literature in three ways. First, it extends research on multimodal literacy by focusing on younger EFL learners, offering insights into how agency develops in primary classrooms. Second, it empirically demonstrates how comprehension gains are linked to agentive engagement with multimodal resources, providing an explanatory account rather than merely reporting outcomes. Third, it employs a mixed-methods design to bridge the gap between quantitative and qualitative approaches, producing findings that are both measurable and richly contextualized. In doing so, it not only advances scholarly understanding of multimodal literacy and epistemic agency but also offers practical implications for pedagogy and policy.

## METHODS

This study employed an explanatory sequential mixed-methods design (Creswell & Plano Clark, 2023) to examine the impact of multimodal reading on EFL learners' comprehension and epistemic agency. This design was chosen because it enabled the researcher to first identify measurable learning gains through quantitative analysis, and then explain these gains by exploring learners' lived experiences through qualitative inquiry. The study was theoretically grounded in multimodality (Jewitt, 2014), which views meaning-making as distributed across linguistic, visual, audio, gestural, and spatial modes, and epistemic agency (Biesta, 2009; Mercer, 2019), which emphasizes learners' initiative, decision-making, and responsibility in constructing knowledge.

### *Phase 1: Quantitative*

A one-group pretest–posttest design was implemented to measure changes in reading comprehension before and after the multimodal reading intervention. The participants were 30 fifth-grade students from State Islamic Elementary School (MIN) 1 Tulang Bawang Barat, Lampung Province recruited using total population sampling. The reading comprehension test, adapted from the national curriculum standards, consisted of 20 multiple-choice items covering literal, inferential, and evaluative comprehension. Content validity was established through expert judgment by two EFL specialists, and reliability analysis yielded a Cronbach's alpha of 0.87, confirming internal consistency.

The intervention was conducted over four weeks (two sessions per week, 90 minutes per session). Students engaged in multimodal reading activities that integrated linguistic texts, visual representations (images, diagrams), audio narration, gestural enactments, and spatial layouts (page design and digital navigation). These activities encouraged learners not only to interpret texts across modes but also to exercise epistemic agency by making predictions, justifying interpretations, and negotiating meaning collaboratively.

### *Phase 2: Qualitative*

In the second phase, qualitative data were collected to explain the quantitative results. Semi-structured interviews were conducted with six students, selected using maximum

variation sampling to represent different levels of comprehension improvement and engagement. The interviews explored learners' experiences of multimodal reading, their strategies for navigating different modes, and their sense of agency in constructing meaning. For the qualitative phase, trustworthiness was enhanced through triangulation, member checking, and peer debriefing, ensuring credibility and interpretive depth.

Data were transcribed and analyzed thematically following Braun & Clarke (2006), combining inductive and deductive coding. Deductive codes were informed by the theoretical constructs of multimodality (Jewitt, 2014) and epistemic agency (Biesta et al., 2015), while inductive coding allowed new insights to emerge from students' accounts.

### *Phase 3: Integrative Phase*

In the final stage, quantitative and qualitative findings were integrated using a joint display (Creswell & Plano Clark, 2023). Quantitative results (pretest–posttest gains) informed the selection of interview participants, while the joint display juxtaposed statistical outcomes with thematic insights (e.g., initiative, reflection, collaboration). This approach enabled direct alignment between what improved and how learners' epistemic agency explained those gains.

### *Data Analysis*

Quantitative data were analyzed using descriptive statistics and paired-sample t-tests to compare pretest and posttest scores (Wallen & Fraenkel, 2013). Qualitative data were analyzed thematically to identify recurring patterns related to learners' multimodal meaning-making and epistemic agency. Integration of the two strands was achieved by comparing numerical trends with thematic insights, thereby providing a comprehensive explanation of how and why multimodal reading influenced learners' comprehension.

## **RESULTS AND DISCUSSION**

### **Quantitative Findings**

The first phase of the study examined whether multimodal reading significantly improved learners' comprehension. Descriptive statistics were calculated to capture the overall distribution of scores, followed by a paired-sample t-test to determine the significance of pretest–posttest differences. The analysis provided evidence not only of measurable improvement but also of the magnitude of the intervention's impact on students' reading performance.

Table 1. Descriptive Statistics of Reading Comprehension Scores (N = 30)

<b>Test</b>	<b>Mean (M)</b>	<b>Standard Deviation (SD)</b>	<b>Minimum</b>	<b>Maximum</b>
Pretest	61.40	8.25	45	75
Posttest	74.30	7.80	60	88

Table 1 presents the descriptive statistics of students' reading comprehension scores before and after the multimodal reading intervention. The pretest mean score was 61.40 (SD = 8.25), indicating that students initially demonstrated a moderate level of comprehension with considerable variation among individuals. After the intervention, the mean posttest score rose to 74.30 (SD = 7.80). The increase of nearly 13 points suggests that students benefited from engaging with multimodal texts, which provided them with additional scaffolds for understanding. The minimum score also improved from 45 on the pretest to 60 on the posttest,

showing that even lower-achieving students made progress. Meanwhile, the maximum score increased from 75 to 88, reflecting the potential of multimodality to push high achievers further. Overall, the descriptive results indicate consistent improvement across the class.

Table 2. Paired-Sample t-test Results for Pretest and Posttest Scores

Pair	Mean Difference	t (df = 29)	p-value	Cohen's d
Pre-Post Test	12.90	8.52	< .001	1.56

Table 2 summarizes the results of the paired-sample t-test conducted to determine whether the observed improvements were statistically significant. The analysis revealed a mean difference of 12.90 points between pretest and posttest scores. This difference was highly significant,  $t (29) = 8.52$ ,  $p < .001$ , providing strong evidence that the intervention effectively enhanced learners' comprehension. The calculated effect size, Cohen's  $d = 1.56$ , indicates a large effect according to Cohen's (1988) benchmark, suggesting that the observed gains were not only statistically significant but also educationally meaningful. This implies that multimodal reading activities had a substantial impact on learners' ability to construct meaning across modes.

The quantitative results demonstrate that multimodal reading yielded both statistically significant and practically meaningful improvements in learners' comprehension. The nearly 13-point gain, coupled with a large effect size ( $d = 1.56$ ), confirms that the intervention had more than a modest impact and was effective across ability levels, as reflected in the raised minimum and maximum scores. These findings align with previous research showing that multimodal scaffolds expand access to meaning and reduce achievement. Notably, the reduction in score variability (SD decreased from 8.25 to 7.80) suggests that the intervention may have narrowed performance disparities within the class. However, as a one-group pretest–posttest design was employed, improvements cannot be attributed exclusively to multimodal instruction without considering possible external influences. Despite this limitation, the magnitude and consistency of gains indicate that the intervention substantially contributed to enhancing comprehension, thereby justifying further exploration through qualitative analysis.

## Qualitative Findings

The second phase explored how learners exercised epistemic agency while engaging with multimodal texts. Semi-structured interviews with six purposively selected students provided insights into their strategies, reflections, and collaborative practices. Thematic analysis revealed recurring patterns that explained the quantitative gains, highlighting how students actively initiated, justified, and regulated their meaning-making across modes.

Table 3. Themes of Learners' Epistemic Agency in Multimodal Reading

Theme	Subtheme / Indicator	Representative Quotes
Taking Initiative	Asking questions, choosing strategies, exploring beyond text	<ul style="list-style-type: none"> <li>“When I saw the picture, I wanted to ask my teacher what it meant” (S1).</li> <li>“I tried to guess the story before reading because the picture gave</li> </ul>

		me ideas" (S3).
Decision-Making Across Modes	Selecting relevant images, prioritizing information, linking visual and text	<ul style="list-style-type: none"> <li>• <i>"I focused on the diagram first because it helped me understand the story" (S2).</i></li> <li>• <i>"The words were long, so I looked at the picture first to know the main idea" (S4).</i></li> </ul>
Justification of Understanding	Explaining reasoning, defending interpretation, using multimodal evidence	<ul style="list-style-type: none"> <li>• <i>"I think my answer is correct because both the picture and the words say the same" (S5).</i></li> <li>• <i>"The gesture in the video shows sadness, that is why I chose that answer" (S1).</i></li> </ul>
Reflection and Self-Regulation	Recognizing difficulties, adjusting strategies, monitoring progress	<ul style="list-style-type: none"> <li>• <i>"At first I only saw the picture, but then I read carefully to check my guess" (S3).</i></li> <li>• <i>"I changed my answer after I saw the diagram matched the text" (S6).</i></li> </ul>
Collaboration and Dialogue	Sharing perspectives, negotiating meaning, building on peers' ideas	<ul style="list-style-type: none"> <li>• <i>"We discussed the gestures in the video and agreed the boy was angry" (S2).</i></li> <li>• <i>"My friend told me a different idea, and I added it to my answer" (S4).</i></li> </ul>

The qualitative findings provided rich insights into how learners exercised epistemic agency while engaging with multimodal texts. A recurring theme was taking initiative, where students demonstrated curiosity and self-directed exploration. For example, S1 described, *"When I saw the picture, I wanted to ask my teacher what it meant"* (S1, Interview), while S3 explained, *"I tried to guess the story before reading because the picture gave me ideas"* (S3, Interview). These actions show that students were not passive recipients but actively initiated meaning-making processes.

Another prominent theme was decision-making across modes, reflecting learners' strategic selection and prioritization of multimodal resources. S2 noted, *"I focused on the diagram first because it helped me understand the story"* (S2, Interview), whereas S4 explained, *"The words were long, so I looked at the picture first to know the main idea"* (S4, Interview). This indicates that learners exercised agency by deliberately navigating different modes to support comprehension.

The theme of justification of understanding highlighted students' ability to defend their interpretations with multimodal evidence. For example, S5 argued, *"I think my answer is correct because both the picture and the words say the same"* (S5, Interview), while S1 emphasized, *"The gesture in the video shows sadness, that is why I chose that answer"* (S1,

Interview). Such reasoning demonstrates epistemic agency in the form of critical evaluation and evidence-based justification.

Learners also engaged in reflection and self-regulation, adjusting strategies when faced with difficulties. As S3 reflected, “*At first I only saw the picture, but then I read carefully to check my guess*” (S3, Interview), while S6 admitted, “*I changed my answer after I saw the diagram matched the text*” (S6, Interview). These reflections reveal metacognitive awareness and a willingness to refine strategies in response to challenges.

Finally, collaboration and dialogue emerged as an essential dimension of epistemic agency. S2 explained, “*We discussed the gestures in the video and agreed the boy was angry*” (S2, Interview), and S4 described, “*My friend told me a different idea, and I added it to my answer*” (S4, Interview). These collaborative interactions indicate that agency was not only individual but also socially distributed, reinforcing the importance of dialogic engagement in multimodal literacy practices.

Finally, these themes show that multimodal reading environments provided opportunities for learners to act agentively by initiating inquiry, making decisions across semiotic modes, justifying their reasoning, reflecting on their strategies, and co-constructing meaning with peers. Such findings align with the theoretical lens that positions epistemic agency as central to knowledge construction, and multimodality as the affordance that enables learners to navigate and integrate diverse meaning-making resources.

### Integrative Findings

The final phase merged the quantitative and qualitative strands to provide a comprehensive explanation of how multimodal reading shaped learners’ comprehension. Using a joint display, statistical improvements were aligned with thematic insights, allowing the study to connect the what of score gains with the how and why of learners’ epistemic agency. This integration demonstrated that measurable improvements in comprehension were underpinned by students’ initiative, decision-making, justification, reflection, and collaboration across modes.

Table 4. Integration of Quantitative and Qualitative Findings

Research Questions	Quantitative Findings	Qualitative Findings	Integrated Interpretation
RQ1: To what extent does multimodal reading improve learners’ comprehension?	Learners’ mean comprehension scores increased significantly from 61.40 (SD = 8.25) in the pretest to 74.30 (SD = 7.80) in the posttest. The gain of 12.90 points was statistically significant, $t(29) = 8.52, p < .001$ , with	Students reported that multimodal resources (visuals, spatial layouts, gestures) supported comprehension by offering alternative entry points to meaning. They described strategies such as predicting from images, cross-	Quantitative improvement is explained by learners’ strategic use of multimodal resources. Engagement with multiple modes enabled students to construct meaning more effectively, thus

	a large effect size ( $d = 1.56$ ).	checking across modes, and verifying answers collaboratively (e.g., <i>"I looked at the diagram first to know the story"</i> – S2, Interview).	driving measurable comprehension gains.
RQ2: How do learners exercise epistemic agency in multimodal reading?	Agency was not directly captured by test scores, but informed the interpretation of comprehension gains.	Five themes of epistemic agency emerged: (1) Taking initiative, (2) Decision-making across modes, (3) Justification of understanding, (4) Reflection and self-regulation, and (5) Collaboration and dialogue. Representative quotes show learners actively directing their own meaning-making process.	The comprehension gains (RQ1) are underpinned by epistemic agency (RQ2). Students exercised agency by initiating strategies, reflecting on choices, and collaborating, which amplified the benefits of multimodal reading.

The integration of quantitative and qualitative results demonstrates a coherent picture of how multimodal reading shaped learners' comprehension and epistemic agency. The quantitative phase (RQ1) showed a statistically significant improvement in comprehension, with learners' mean scores rising from 61.40 to 74.30. The large effect size ( $d = 1.56$ ) indicates that the intervention was not only effective but also had a substantial educational impact.

The qualitative phase (RQ2) adds explanatory depth to these gains by showing that learners did not passively receive information; instead, they exercised epistemic agency through multimodal resources. They initiated strategies such as predicting meaning from images, prioritizing diagrams or layouts before reading the text, and cross-checking answers across different modes. In doing so, students demonstrated agency by making decisions, justifying their interpretations, and regulating their comprehension processes.

This alignment between numerical improvement and thematic insights underscores that the gains in comprehension were driven by learners' active agency. Without the exercise of initiative, reflection, and collaboration, the multimodal materials alone may not have produced such substantial results. Thus, the findings confirm that multimodal reading not only provides multiple channels of representation (Jewitt, 2014) but also creates spaces for learners

to assert control over knowledge construction, consistent with epistemic agency frameworks (Biesta, 2009; Mercer, 2019).

In summary, the explanatory sequential design reveals that the “what” of improved comprehension (Phase 1) is explained by the “how” of learner agency in multimodal interaction (Phase 2). Together, these findings highlight the pedagogical value of multimodal resources when they are coupled with opportunities for learners to exercise epistemic agency.

## **Discussion**

The significant quantitative gains in reading comprehension suggest that multimodal resources provided powerful scaffolds for meaning-making. This supports the growing body of evidence that multimodality enhances comprehension by integrating visual, spatial, and gestural cues with linguistic input (Daulay & Dewi, 2025; Hardison & Pennington, 2021; Zhang et al., 2022). Unlike earlier literacy approaches that relied predominantly on linear text, multimodal pedagogies acknowledge the semiotic richness of texts and allow learners to mobilize diverse cues for interpretation (Jewitt, 2014; Mills, 2016). The results of this study extend these insights by showing that such integration was not incidental but actively driven by learners’ agentive engagement.

The qualitative findings further illuminated the mechanisms behind these gains: learners took initiative, justified their answers, and reflected on their strategies across modes. These behaviors exemplify epistemic agency, positioning students as decision-makers in the learning process (Biesta et al., 2015; Van de Pol et al., 2019; Vaughn et al., 2020). Previous research shows that when learners are invited to exercise epistemic authority, they become more engaged and construct knowledge collaboratively (Eriksson & Lindberg, 2016; Yang, 2019; Zhang et al., 2022). The present findings converge with this literature by demonstrating how agency manifested concretely in multimodal tasks such as interpreting diagrams, gestures, or spatial layouts.

Importantly, the integration of agency with multimodality highlights that comprehension improvement is not attributable to semiotic resources alone but to how learners activated them. This resonates with the view that multimodal texts afford possibilities, but learners’ choices and reflections determine whether these affordances become pedagogically meaningful (Martin, 2020; Zhang et al., 2022). In our study, students did not simply notice images or gestures; they strategically used them as evidence, a process consistent with reflective agency and knowledge-building pedagogy (Jewitt, 2014; Rappa & Tang, 2018).

These findings also align with research on differentiated instruction and universal design for learning, which argue that multiple pathways to meaning support a wider range of learners (Mayer, 2014; Wan, 2017). The fact that lower-achieving students particularly benefited from multimodal activities echoes Daulay & Dewi’s (2025) conclusion that multimodal pedagogy narrows achievement gaps. Similar studies have shown that visual and spatial resources are especially valuable for learners who struggle with linguistic decoding, allowing them to access higher-order comprehension (January & Nima, 2018; Santoso & Mangkuluhur, 2024; White, 2024).

At the same time, the study underscores the relational nature of epistemic agency. Collaborative dialogue emerged as a key practice through which students co-constructed

meaning, consistent with sociocultural theories of learning (Eriksson & Lindberg, 2016; Martin, 2020). Prior research has shown that dialogic engagement fosters accountability and deepens comprehension in literacy classrooms (Vaughn et al., 2020; Zhang et al., 2022). In our data, students often refined their interpretations by negotiating gestures, images, and textual evidence with peers, illustrating that agency is socially distributed rather than purely individual.

Nevertheless, multimodal pedagogy also carries challenges. Cognitive load theory reminds us that too many modes can overwhelm learners if not scaffolded carefully (Mayer, 2014; Weninger, 2023; White, 2024). Some students initially focused only on visuals and had to adjust by cross-checking with text, a finding consistent with studies highlighting the need for guided structuration in multimodal tasks (Mills, 2016; Zhang et al., 2022). This suggests that teachers must balance freedom for agency with explicit scaffolding to ensure that multimodal affordances are transformed into effective strategies rather than sources of confusion.

In summary, the findings affirm that multimodal literacy cannot be separated from epistemic agency. Learners improved comprehension not simply because they were exposed to multiple modes, but because they actively navigated, selected, and justified meaning across those modes. This echoes global calls for 21st-century literacies that integrate criticality, collaboration, and self-direction into reading pedagogy (Cope & Kalantzis, 2020; Jewitt, 2014). By situating learners as epistemic agents, multimodal reading equips them with both measurable literacy skills and broader competencies for knowledge construction in complex communicative environments.

## CONCLUSION

This study highlights that fostering multimodal literacy is most effective when it is understood as an agentive process of knowledge construction rather than mere exposure to multiple modes. When learners are positioned as epistemic agents able to choose, justify, and reflect on their meaning-making strategies multimodal reading becomes a space for developing critical thinking, self-regulation, and collaborative learning. Such a perspective shifts literacy pedagogy toward empowering students to navigate complex semiotic environments and construct understanding actively. These insights call for instructional designs that balance scaffolding with learner autonomy, curriculum policies that recognize multimodal competence as part of core literacy development, and further research exploring how agency can be nurtured across diverse contexts and digital platforms.

## REFERENCES

AlAli, R., & Wardat, Y. (2024). Low PISA performance students: Factors, perceptions, and improvement strategies. *International Journal of Religion*, 5(9), 334–348. <https://doi.org/10.61707/nve8gj33>

Biesta, G., Priestley, M., & Robinson, S. (2015). The role of beliefs in teacher agency. *Teachers and Teaching*, 21(6), 624–640. <https://doi.org/10.1080/13540602.2015.1044325>

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

Cope, B., & Kalantzis, M. (2020). *Making Sense: Reference, Agency, and Structure in a Grammar of Multimodal Meaning*. <https://doi.org/10.1017/9781316459645>

Creswell, J. W., & Plano Clark, V. L. (2023). Revisiting mixed methods research designs twenty years later. *Handbook of Mixed Methods Research Designs*, 1(1), 21–36. <http://dx.doi.org/10.4135/9781529614572.n6>

Daulay, E., & Dewi, U. (2025). Reading Comprehension through Multimodal Literacy Approach for English Education. *ETERNAL (English Teaching Journal)*, 16(2), 397–404. <https://doi.org/10.26877/a5zsp527>

Eriksson, I., & Lindberg, V. (2016). Enriching ‘learning activity’ with “epistemic practices”—enhancing students’ epistemic agency and authority. *Nordic Journal of Studies in Educational Policy*, 2016(1), 32432. <https://doi.org/10.3402/nstep.v2.32432>

Fälth, L., Brkovic, I., Kerestes, G., Svensson, I., Hjelmquist, E., & Tjus, T. (2023). The effects of a multimodal intervention on the reading skills of struggling students: An exploration across countries. *Reading Psychology*, 44(3), 225–241. <https://doi.org/10.1080/02702711.2022.2141399>

Hardison, D. M., & Pennington, M. C. (2021). Multimodal second-language communication: Research findings and pedagogical implications. *Relc Journal*, 52(1), 62–76. <https://doi.org/10.1177/0033688220966635>

January, R., & Nima, H. N. A. (2018). Energizing Students’ Reading Comprehension through Multimodal Texts. *International Journal of Language Education*, 2(2), 14–22. <https://doi.org/10.26858/ijole.v2i2.4347>

Jewitt, C. (2014). 12. Multimodal approaches. *Interactions, Images and Texts: A Reader in Multimodality*, 11, 127. <http://dx.doi.org/10.1515/9781614511175.127>

Kim, J., Yu, S., Lee, S.-S., & Detrick, R. (2025). Students’ prompt patterns and its effects in AI-assisted academic writing: Focusing on students’ level of AI literacy. *Journal of Research on Technology in Education*, 1–18. <http://dx.doi.org/10.1080/15391523.2025.2456043>

Mahapoonyanont, N., & Songsang, N. (2024). Policy learning and adaptation: Lessons from PISA for educational reform worldwide. *International Journal of Stress Management*, 31(4), 26–52. <http://dx.doi.org/10.5281/zenodo.14179522>

Martin, D. (2020). Providing Students with multimodal feedback experiences. *Journal of Curriculum, Teaching, Learning and Leadership in Education*, 5(1), 16. <https://doi.org/10.32873/uno.dc.ctlle.05.01.1082>

Mayer, R. E. (2014). Based principles for designing multimedia instruction. *Copyright and Other Legal Notices*, 59. <http://dx.doi.org/10.1037/0003-066X.63.8.760>

Mills, C. B. (2016). *The Effect of Digital Media on Emergent Literacy Skills: A Systematic Review*. <http://dx.doi.org/10.13023/ETD.2016.071>

Norman, A. (2023). Educational technology for reading instruction in developing countries: A systematic literature review. *Review of Education*, 11(3), e3423. <https://doi.org/10.1002/rev3.3423>

Rappa, N. A., & Tang, K.-S. (2018). Integrating disciplinary-specific genre structure in discourse strategies to support disciplinary literacy. *Linguistics and Education*, 43, 1–12. <http://dx.doi.org/10.1016/j.linged.2017.12.003>

Rohi, M. P., & Nurhayati, L. (2024). Multimodal learning strategies in secondary EFL

education: Insights from teachers. *Voices of English Language Education Society*, 8(2). <https://doi.org/10.29408/veles.v8i2.26546>

Santoso, W., & Mangkuluhur, B. D. (2024). Pre-Service English Teachers' Perceptions Towards The Use of Dialogical Learning as Reflective Practice: A Case Study. *SALEE: Study of Applied Linguistics and English Education*, 5(2), 493–507. <http://dx.doi.org/10.35961/salee.v5i2.1401>

Van de Pol, J., Mercer, N., & Volman, M. (2019). Scaffolding student understanding in small-group work: Students' uptake of teacher support in subsequent small-group interaction. *Journal of the Learning Sciences*, 28(2), 206–239. <http://dx.doi.org/10.1080/10508406.2018.1522258>

Vaughn, M., Jang, B. G., Sotirovska, V., & Cooper-Novack, G. (2020). Student agency in literacy: A systematic review of the literature. *Reading Psychology*, 41(7), 712–734. <http://dx.doi.org/10.1080/02702711.2020.1783142>

Wallen, N. E., & Fraenkel, J. R. (2013). Experimental Research. In *Educational Research* (pp. 279–306). Routledge. <https://doi.org/10.4324/9781410601001>

Wan, S. W.-Y. (2017). Differentiated instruction: Are Hong Kong in-service teachers ready? *Teachers and Teaching*, 23(3), 284–311. <http://dx.doi.org/10.1080/13540602.2016.1204289>

Weninger, C. (2023). Digital literacy as ideological practice. *Elt Journal*, 77(2), 197–206. <http://dx.doi.org/10.1093/elt/ccad001>

White, J. P. (2024). Unlocking Potential With Multimodal Learning and Assessment: A Discussion of the Barriers and Benefits. *GILE Journal of Skills Development*, 4(2), 106–111. <http://dx.doi.org/10.52398/gjsd.2024.v4.i2.pp106-111>

Yang, Y. (2019). Reflective assessment for epistemic agency of academically low-achieving students. *Journal of Computer Assisted Learning*, 35(4), 459–475. <http://dx.doi.org/10.1111/jcal.12343>

Yi, Q., Dong, Z., & Qiao, H. (2024). Enhancing EFL listening and speaking skills: Strategies and practice for implementing multimedia and multi-modal approaches. *Journal of Humanities, Arts and Social Science*, 8(9). <http://dx.doi.org/10.26855/jhass.2024.09.032>

Yi, Q., Zhao, K., & Qiao, H. (2024). The Dynamics of Multi-modal Practice and Pedagogy in EFL Instruction: A Theoretical Exploration. *The Educational Review, USA*, 8(9). <http://dx.doi.org/10.26855/er.2024.09.011>

Zhang, J., Tian, Y., Yuan, G., & Tao, D. (2022). Epistemic agency for costructuring expansive knowledge-building practices. *Science Education*, 106(4), 890–923. <https://doi.org/10.1002/sce.21717>