Proceedings of Forum for University Scholars in Interdisciplinary Opportunities and Networking
Graduate School, Universitas Terbuka
https://conference.ut.ac.id/index.php/fusion

Exploring Students' Engagement with Multimodal Texts in English Learning Activities

Khusnul Harsul Lisan[a*], Adhan Kholis[b] , Ulul Azmi[c]
[a*]Universitas Nahdlatul Ulama Yogyakarta, Yogyakarta, Indonesia, khusnulharsullsn@gmail.com
[b]Universitas Nahdlatul Ulama Yogyakarta, Yogyakarta, Indonesia, adhan@unu-jogja.ac.id
[c]Universitas Nahdlatul Ulama Yogyakarta, Yogyakarta, Indonesia, luluk@unu-jogja.ac.id

*Correspondence: khusnulharsullsn@gmail.com

Abstract
This study explores students' engagement with multimodal texts in the English learning activities at a senior high school in Yogyakarta. As digital technologies continue to reshape the educational landscape, students are increasingly immersed in diverse forms of texts that integrate written, visual, audio, and interactive elements during the subject. These multimodal texts require a broader set of literacy skills that go beyond traditional reading and writing. While such texts are becoming more common in 21st-century classrooms, especially in language learning environments, the ways in which students engage with them in practice remain insufficiently examined, particularly within the Indonesian English as foreign language context. This qualitative case study involved some classes of Grade XII students of Ali Maksum Islamic Senior High School and the English teacher as participants. The data were collected through classroom observations, student interviews, and document analysis of teaching materials. The findings revealed that students engage with multimodal texts in three key dimensions: cognitive (understanding content), behavioural (participation and task completion), and affective (interest and motivation). Visual, audio, and video-based texts and other interactive elements were found to enhance student interest, engagement, literacy and comprehension. The study highlights the importance of integrating multimodal literacy practices into English instruction to support students' active engagement and improve learning outcomes in 21st-century classrooms along with the required skills.

Keywords:
English subject;
Students' engagement;
Multimodal texts

1. Introduction

Literacy in the 21st century has expanded far beyond the traditional ability to read and write printed texts. In contemporary education, communication is increasingly mediated by digital technologies, multimedia platforms, and diverse semiotic resources. Within this context, the concept of multiliteracies (New London Group, 1996; Cope & Kalantzis, 2009) emphasizes the capacity of learners to interpret, create, and engage with meaning across multiple modes, including visual, spatial, auditory, and gestural forms. Such an approach recognizes that meaning-making in today's classrooms can no longer be confined to alphabetic texts alone, but must extend to visual design, sound, movement, and digital interactivity. This shift challenges educators to reconceptualize literacy not merely as decoding and encoding words, but as navigating and orchestrating multimodal signs (Jewitt, Bezemer, & O'Halloran, 2016).

Multimodal texts—such as videos, infographics, podcasts, animations, and interactive learning platforms—play a central role in this new literacy landscape (Cope & Kalantzis, 2015). In English language learning, these texts are not merely supplementary but represent innovative ways of

constructing knowledge, scaffolding comprehension, and fostering active learner engagement (Walsh, 2010; Kress & van Leeuwen, 2001). For example, a short video clip or infographic can convey complex ideas more efficiently than linear written passages, while podcasts and interactive applications allow students to revisit materials and develop listening, speaking, and critical literacy skills simultaneously. These diverse resources expand opportunities for differentiated learning and help teachers respond to students' varied literacy repertoires (Serafini, 2012; Hafner, 2014). Recent studies also suggest that multimodal approaches foster collaboration and creativity, both of which are essential skills in 21st-century education (Mills, 2016; Mercer & Dörnyei, 2020).

The rapid advancement of artificial intelligence, particularly in the field of deep learning, has further accelerated the development, circulation, and accessibility of multimodal texts (Li et al., 2022; Xu et al., 2023). Deep learning models which are also introduced by the government in these recent curriculum focuses on how students increase their competence in critical thinking. Cholifatunisa et al. (2025) explain that the development of Merdeka curriculum with deep learning approach emphasize the students competence not only in critical thinking but also solve the real problem. In fact, there are so many tools which is supportive like machine translation, and image-to-text generation, all of which increasingly shape the ways multimodal content is produced and consumed in education. In addition, multimodal learning analytics powered by deep learning can capture patterns of student interaction across text, audio, and video, offering new possibilities for understanding engagement (Borthwick et al., 2021; Luckin, 2022). These innovations highlight that multimodality is not only a pedagogical choice but also a technological reality of the digital age. However, while students—often labeled as digital natives (Prensky, 2001)—are surrounded by these technologies, exposure alone does not guarantee meaningful engagement or effective comprehension (Rowsell & Walsh, 2011; Godwin-Jones, 2021).

For teachers, therefore, the challenge lies in designing and mediating learning experiences that transform students from passive consumers of multimodal texts into critical, active participants. Research has shown that pedagogically guided use of multimodal resources can enhance vocabulary acquisition, strengthen critical thinking, and promote collaboration (Serafini, 2012; Hafner, 2014; Ajayi, 2009). Yet, in practice, teachers often struggle with limited training, resource constraints, and lack of familiarity with multimodal pedagogy (Hapsari, 2021). This gap between theoretical recognition and classroom implementation is particularly visible in contexts such as Indonesia, where curriculum policies like the 2013 Curriculum and the more recent *Kurikulum Merdeka* explicitly encourage the integration of digital and multimodal resources, but classroom practices remain heavily text-based and teacher-centered (Agustina & Cahyono, 2017; Kusumaningrum et al., 2024). Having the multiliteracies activities, the teacher are supposed to get facilitated in delivering the English materials with some variations of multimodal texts. It becomes very essential regarding the input that they provide for the students. The variation of multimodal texts elicit students' motivation towards English materials.

In this regard, it becomes important to investigate not only whether multimodal texts are being used, but also how students actually engage with them. Engagement is a multifaceted construct encompassing students' cognitive investment, behavioral participation, and emotional involvement (Fredricks, Blumenfeld, & Paris, 2004). Cognitive engagement refers to the mental strategies and effort students employ to process information; behavioral engagement encompasses their participation, attentiveness, and task completion; and affective engagement reflects their motivation, interest, and emotional connection to the learning experience. Exploring these dimensions within multimodal learning environments provides valuable insights into how multimodal literacy can be effectively cultivated (Philp & Duchesne, 2016; Mercer & Dörnyei, 2020).

Despite the growing interest in multimodal literacy worldwide, research on student engagement with multimodal texts in Indonesian EFL classrooms remains limited. Previous studies tend to focus on teachers' perceptions (Hapsari, 2021), specific media types such as digital comics or videos (Agustina & Cahyono, 2017), or early education contexts (Insani, Suwarna, & Triyono, 2024). Much less attention has been given to senior high school learners, whose literacy practices are more complex and whose learning outcomes are critical for higher education readiness. Addressing this gap, the present study investigates how Grade XII students at a senior high school in Yogyakarta engage cognitively, behaviorally, and affectively with multimodal texts in English learning activities. By employing a qualitative case study approach, this research aims to illuminate the dynamics of engagement in multimodal learning contexts and to provide pedagogical implications for enhancing English instruction

in 21st-century Indonesian classrooms. Hopefully, this research illuminates the teachers or educators who are willing to maximize the potential and engagement of students towards English materials by using multiliteracy activities.

2. Method

This study employed a qualitative case study design to explore students' engagement with multimodal texts in English learning activities. A case study approach was deemed appropriate because it allows for an in-depth investigation of a bounded system within its real-life educational context (Creswell, 2018; Yin, 2014). The focus of this research was not on generalizing to a large population but on generating rich, contextualized insights into how students interact with multimodal resources in a specific classroom setting. In line with Merriam and Tisdell (2016), the case study design facilitated the holistic investigation of students' cognitive, behavioral, and affective engagement as they encountered multimodal texts during classroom activities.

The study was conducted at Ali Maksum Islamic Senior High School, Yogyakarta, Indonesia, a private boarding school with a strong emphasis on academic achievement and religious values. The school has integrated aspects of the 2013 Curriculum and *Kurikulum Merdeka*, both of which encourage the use of diverse learning resources, including multimodal materials. However, English instruction in the school still relies heavily on traditional print-based textbooks, making it a suitable site for examining how multimodal texts are incorporated and experienced in practice. Participants included one English teacher and three classes of Grade XII students (approximately 90 students in total). The teacher was selected due to her active efforts in integrating videos, audio recordings, and digital materials into English lessons. Students were chosen purposively, as senior high school learners at this level are expected to demonstrate higher-order literacy skills and prepare for university entrance examinations. To protect anonymity, pseudonyms were assigned to all participants in reporting the findings.

Data were collected over a period of two months through three primary techniques: classroom observations, student interviews, and document analysis.

a) Classroom Observations: A total of three classroom were observed, focusing on how multimodal texts were used and how students responded during lessons. Field notes captured instances of cognitive engagement (e.g., questioning, summarizing, connecting ideas), behavioral engagement (e.g., participation, attentiveness, task completion), and affective engagement (e.g., expressions of interest, enthusiasm, frustration). Selected sessions were video-recorded with the teacher's consent to allow for more detailed analysis;
b) Student Interviews: Semi-structured interviews were conducted with twelve students representing varying levels of participation and achievement. Interview questions explored their perceptions of multimodal texts, learning strategies, motivations, and challenges. This method provided deeper insights into students' cognitive and affective experiences that could not be fully observed in classroom settings;
c) Document Analysis: Teaching materials such as lesson plans, PowerPoint slides, worksheets, and digital media resources (e.g., YouTube videos, audio recordings, infographics) were collected and analyzed. This helped identify the types of multimodal texts used and the pedagogical intentions behind their integration.

Triangulating these three data sources strengthened the credibility of the findings and allowed for a more comprehensive picture of student engagement.

Data analysis followed the interactive model of Miles, Huberman, and Saldaña (2014), consisting of data condensation, data display, and conclusion drawing/verification. Observation notes, interview transcripts, and documents were first coded inductively to capture patterns of engagement. Codes were then categorized under the three dimensions of engagement: cognitive, behavioral, and affective (Fredricks et al., 2004). Within each category, themes and sub-themes were identified, such as comprehension strategies, participation in group tasks, or emotional responses to video materials. Displays in the form of matrices and charts were used to organize the data and facilitate interpretation. Throughout the analysis, themes were continuously compared and refined to ensure alignment with the research objectives.

3. Results and Discussion

3.1 Cognitive Engagement

Students demonstrated strong cognitive engagement when interacting with multimodal texts, particularly videos, infographics, and illustrated worksheets. Visual inputs enhanced comprehension and triggered more student-generated questions compared to text-only lessons. For example, during a lesson on report texts, an infographic illustrating the structure ("general classification," "description," and "closing") helped students recall key elements while completing comprehension exercises. Several students explicitly referred back to the infographic when answering questions, showing how visual scaffolds supported memory retention and comprehension. One student shared in an interview:

> *"When the teacher showed us the infographic, I could remember the parts of the report text more easily. I usually forget after reading the book, but with pictures it stays in my mind." (Student Interview, March 2025)*

This finding aligns with Walsh (2010), who stresses that multimodal texts provide crucial scaffolding for second-language learners by reinforcing comprehension through multiple semiotic channels. Similarly, Kress and van Leeuwen's (2001) framework of multimodal discourse explains how learners actively integrate visual and linguistic modes to construct meaning.

However, when multimodal content lacked sufficient scaffolding, comprehension challenges surfaced. For instance, in one activity students were asked to interpret memes in English. While some were able to connect the humor with cultural context, several misinterpreted the intended meaning. One student admitted:

> *"I didn't get the joke in the meme. Maybe because it was about American culture, so I just guessed." (Student Interview, March 2025)*

The teacher also reflected on this challenge:

> *"Students enjoy memes, but without explanation, they sometimes misunderstand the meaning. I need to give cultural context before asking them to interpret." (Teacher Interview, March 2025)*

This highlights the importance of guided analysis, echoing Cope and Kalantzis (2009), who emphasize that exposure to multimodal texts must be accompanied by explicit instruction to develop critical literacy. On the top of that, the student felt enjoyed and excited during the lesson although in certain messages or information, the teacher needed to explain more regarding the contexts. This kind of multimodal sources are very essential to elicit student's critical thinking towards the English lesson.

3.2 Behavioral Engagement

The study also revealed that multimodal activities fostered higher levels of behavioral engagement, including participation, sustained attention, and task completion. Group projects that required students to analyze videos and create digital posters encouraged collaboration and ownership of learning. For example, during a lesson on environmental issues, students were asked to watch a short video and summarize their reflections using Canva or other tools. All groups completed the task on time, with lively discussions observed throughout the process. One student remarked:

> *"I usually feel sleepy in English class, but when we used Canva to make a poster, I really enjoyed it because we could be creative and share my ideas since I like designing somethings through visual tools." (Student Interview, March 2025)*

Observation notes confirmed this behavioral shift. Students who were typically passive during textbook-based exercises became more active in contributing ideas during digital poster creation. The teacher observed similar changes:

> *"When I give them multimodal tasks, even the quiet students try to say something or contribute. It makes the class more alive compared to only using the book." (Teacher Interview, March 2025)*

This resonates with Fredricks et al. (2004), who argue that behavioral engagement is visible in persistence and effort. It also supports Hafner's (2014) findings that multimodal projects foster collaborative learning and sustained attention. Based on the interview, the students said that the good attitude or behaviour shows a positive impact of multimodal texts. It affects to the students' attention which increased significantly. It is known that the characteristics of students in islamic boarding schools are getting easy to fall a slep during the lesson. The multiliteracy pedogogy somehow change gradually how the students behaved during the class where many of them are motivated to give contribution in finishing the task completion in variaty of activities which is in the scope of multimodal texts.

Nevertheless, disparities in access to digital tools emerged as a barrier. Two students admitted difficulty completing digital assignments at home due to unstable internet connections. As one noted:

> *"I like doing digital projects, but at the boarding schools, we don't always have internet, so sometimes my group and I cannot finish the task properly." (Student Interview, March 2025)*

Such challenges reflect the equity concerns raised by Cope and Kalantzis (2015), who caution that multimodal pedagogy must consider infrastructural gaps to avoid reinforcing educational inequalities. The reality is that the students in this boarding school wouldn't be able to use internet or the computer devices all the time. They only manage to deal with such tools in particular times which are provided by the supervisors and teachers in the boarding schools.

3.3 Affective engagement

Affective engagement was evident in students' positive emotional responses to multimodal texts. Most participants described lessons involving videos, songs, and infographics as "fun," "less boring," and "more motivating" compared to traditional textbook-based tasks. Laughter, spontaneous comments, and enthusiastic gestures were frequently observed during video-based lessons, indicating heightened enjoyment and interest. One student explained:

> *"I feel more motivated when the class starts with a video. It makes me curious about the topic, so I want to pay more attention." (Student Interview, March 2025)*

The teacher echoed this observation:

> *"If I start the lesson with a video or song, students are immediately more enthusiastic. Their mood is different, more positive, compared to when I begin with only reading texts." (Teacher Interview, March 2025)*

These findings align with Dörnyei's (2009) emphasis on the motivational dimension of language learning and Fredricks et al.'s (2004) view that affective engagement is critical for sustaining long-term interest. However, affective engagement was not uniform across all students. A few expressed frustration when videos were too fast-paced or when teachers skipped explanations of visual content. As one student put it:

> *"Sometimes the video is too fast. I cannot catch the English words, and then I feel lost." (Student Interview, March 2025)*

This underscores the importance of instructional pacing and scaffolding in sustaining positive emotions, as suggested by Rowsell and Walsh (2011). In addition, the teacher should be able to perceive the students' interest since not all of the students have same characteristics. This means that there should be an analysis done by the teacher to accommodate their interests to be put in the English learning materials through variaty of multimodal texts.

Table 1
Summary of Findings

Engagement Type	Indicators	Evidence Sources
Cognitive	Improved comprehension with visuals; deeper questioning; challenges with cultural references	Observations, Student Interviews, Teacher Interviews, Documents
Behavioral	Increased participation, collaboration, timely task completion; digital access barriers	Observations, Group Work, Interviews
Affective	Positive attitudes, enthusiasm, enjoyment; occasional frustration with fast-paced input	Interviews, Observations

3.4 Integrating the Dimensions

The findings suggest that cognitive, behavioral, and affective engagement are deeply interconnected. Students who comprehended content more effectively through multimodal texts tended to show greater participation and more positive attitudes. Conversely, lack of comprehension sometimes led to disengagement, both behaviorally and emotionally. This interplay confirms Fredricks et al.'s (2004) multidimensional framework of engagement.

Moreover, the evidence underscores the importance of teacher mediation and equitable access**.** While multimodal texts hold significant potential to enhance learning, their effectiveness relies on teachers' ability to scaffold critical engagement and ensure that all students—regardless of background—can participate meaningfully. In light of the rapid growth of digital and AI-driven learning

resources, Indonesian EFL classrooms must be equipped not only with infrastructure but also with pedagogical strategies that maximize the benefits of multimodal literacy.

3.5 Pedagogical Implications

While this study identified several challenges in implementing multimodal learning—such as limited internet access, unequal digital literacy, and restricted technological facilities—it also highlights feasible pedagogical strategies for ensuring effective and equitable implementation within resource-constrained contexts. These implications align with prior studies emphasizing that multimodal pedagogy requires both *technological readiness* and *pedagogical innovation* (Cope & Kalantzis, 2015; Jewitt et al., 2016).

First, teachers can adopt a blended multimodal approach that integrates both offline and online resources. Offline materials such as printed infographics, locally stored videos, and audio files provide multimodal input without relying heavily on internet infrastructure. When connectivity allows, online tools like Padlet, YouTube, or Canva can be selectively employed to extend learning opportunities. This approach aligns with the concept of *multiliteracies pedagogy*, which encourages the flexible use of multiple modes to enhance accessibility and engagement (New London Group, 1996; Mills, 2016).

Second, peer-assisted learning can be an effective strategy for addressing disparities in students' digital competence. Collaborative learning promotes shared expertise and reduces individual gaps in technology use. Research by Mercer and Dörnyei (2020) and Hafner (2014) indicates that multimodal group projects encourage cooperative engagement, foster social responsibility, and help students co-construct meaning through negotiation and collaboration.

Third, teacher scaffolding plays a critical role in maximizing multimodal engagement. Teachers should provide pre-task guidance such as vocabulary pre-teaching, contextual explanation, or comprehension questions before presenting multimodal materials. This resonates with Serafini's (2012) and Walsh's (2010) arguments that explicit scaffolding supports learners in interpreting multimodal signs critically, preventing surface-level engagement and promoting deep comprehension.

Fourth, sustainable multimodal learning requires institutional and policy-level support. Schools should provide continuous professional development on multimodal pedagogy and equip classrooms with basic technological tools such as projectors, speakers, or local storage. As Cope and Kalantzis (2009) emphasize, effective multimodal teaching depends not only on access to technology but also on teachers' ability to design transformative learning experiences that integrate digital and traditional literacies.

Finally, multimodal learning should be embedded gradually within the existing curriculum rather than treated as an isolated innovation. By aligning multimodal tasks with curricular objectives, assessment criteria, and classroom realities, teachers can ensure consistency and sustainability. This view echoes Hafner (2014) and Agustina & Cahyono (2017), who note that successful multimodal implementation in EFL classrooms depends on coherence between pedagogy, resources, and context. Taken together, these pedagogical implications demonstrate that multimodal learning can be both equitable and effective when teachers adopt creative, context-aware, and research-informed approaches to instruction.

4. Conclusion

This study explored how senior high school students in Yogyakarta engaged with multimodal texts in English learning. The findings revealed that students' engagement occurred across three interconnected dimensions—cognitive, behavioral, and affective—which together shaped their overall learning experience. Multimodal resources such as videos, infographics, and audio materials supported comprehension, fostered collaboration, and enhanced students' motivation, yet several challenges emerged, including limited digital access and cultural misunderstandings.

Theoretically, this research extends Fredricks et al.'s (2004) framework of engagement by demonstrating how these dimensions interact within a multimodal and EFL context. It highlights that engagement with multimodal texts is not merely about exposure to diverse media but about the quality of instructional mediation that encourages critical thinking and reflection. This adds a new layer to existing studies by showing how multimodal literacy can strengthen cognitive depth and affective involvement simultaneously. Practically, the study suggests several pedagogical implications. Teachers

should design multimodal activities that integrate visual and textual analysis, foster collaborative work, and scaffold cultural understanding. Schools and policymakers should support teachers through professional development in multimodal pedagogy and improve facilities to ensure equitable access to digital tools. These actions will enable students to benefit from multimodal learning regardless of infrastructure constraints.

Future research may compare different educational settings—urban and rural, public and private—to understand how contextual factors influence multimodal engagement. Longitudinal studies could also examine how consistent exposure to multimodal texts influences students' digital literacy and language proficiency over time. Through these directions, this study contributes both conceptually and practically to enhancing English learning in the era of 21st-century literacies.

5. References

Agustina, D., & Cahyono, B. Y. (2017). Multimodal texts in Indonesian EFL classrooms: Teachers' practices and challenges. *Indonesian Journal of Applied Linguistics, 7*(2), 283–291. https://doi.org/10.17509/ijal.v7i2.8350

Ajayi, L. (2009). English as a second language learners' exploration of multimodal texts in a junior high school. *Journal of Adolescent & Adult Literacy, 52*(7), 585–595. https://doi.org/10.1598/JAAL.52.7.4

Borthwick, A., Hansen, R., Gray, L., & Ziemann, I. (2021). Deep learning for multimodal learning analytics: Potentials and challenges. *Computers & Education: Artificial Intelligence, 2*, 100027. https://doi.org/10.1016/j.caeai.2021.100027

Cholifatunisa, N., Aulia, R., Marlina, D., & Iskandar, A. (2025). Pengembangan Kurikulum Merdeka dengan pendekatan deep learning dalam meningkatkan kompetensi siswa sekolah dasar. *Jurnal Pendidikan Dasar*, 12(1), 45–56. https://ejournal.upi.edu/index.php/jppd/article/view/84240

Cope, B., & Kalantzis, M. (2009). "Multiliteracies": New literacies, new learning. *Pedagogies: An International Journal, 4*(3), 164–195. https://doi.org/10.1080/15544800903076044

Cope, B., & Kalantzis, M. (2015). *The things you do to know: An introduction to the pedagogy of multiliteracies*. In A. Healy (Ed.), *Multiliteracies and diversity in education: New pedagogies for expanding landscapes* (pp. 1–36). Oxford University Press.

Dörnyei, Z. (2009). *The psychology of second language acquisition*. Oxford University Press.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 74*(1), 59–109. https://doi.org/10.3102/00346543074001059

Godwin-Jones, R. (2021). Emerging technologies: Digital literacies revisited. *Language Learning & Technology, 25*(3), 1–7. https://doi.org/10.10125/73410

Hafner, C. A. (2014). Embedding digital literacies in English language teaching: Students' digital video projects as multimodal ensembles. *TESOL Quarterly, 48*(4), 655–685. https://doi.org/10.1002/tesq.138

Hapsari, I. I. (2021). Teachers' perceptions and challenges in implementing multimodal literacy in Indonesian classrooms. *Journal of English Language Teaching and Linguistics, 6*(2), 349–361. https://doi.org/10.21462/jeltl.v6i2.542

Insani, N. H., Suwarna, & Triyono, S. (2024). Effect of multimodal literacy on reading ability of Indonesian Javanese learners. *Issues in Language Studies, 13*(1), 23–38. https://publisher.unimas.my/ojs/index.php/ILS/article/view/6472

Jewitt, C., Bezemer, J., & O'Halloran, K. (2016). *Introducing multimodality*. Routledge.

Kress, G., & van Leeuwen, T. (2001). *Multimodal discourse: The modes and media of contemporary communication*. Arnold.

Kusumaningrum, S. R., Utami, T. S., & Wulandari, S. (2024). Multimodal texts implementation in Indonesian primary and secondary EFL classrooms: Prospects and challenges. *Jurnal Pendidikan Progresif, 14*(1), 1–15. https://jurnal.fkip.unila.ac.id/index.php/jpp/article/view/31676

Li, X., Chen, Y., & Zhang, H. (2022). Multimodal deep learning for educational applications: A systematic review. *Educational Technology Research and Development, 70*(4), 1503–1525. https://doi.org/10.1007/s11423-022-10123-y

Luckin, R. (2022). Machine learning and artificial intelligence in education: Promises and pitfalls. *British Journal of Educational Technology, 53*(6), 1341–1356. https://doi.org/10.1111/bjet.13217

Mercer, S., & Dörnyei, Z. (2020). *Engaging language learners in contemporary classrooms*. Cambridge University Press.

Mills, K. A. (2016). Literacy theories for the digital age: Social, critical, multimodal, spatial, material and sensory lenses. *Multimodal Literacies and Emerging Genres, 32*(2), 214–229. https://doi.org/10.1080/15427587.2016.1141143

New London Group. (1996). A pedagogy of multiliteracies: Designing social futures. *Harvard Educational Review, 66*(1), 60–92. https://doi.org/10.17763/haer.66.1.17370n67v22j160u

Philp, J., & Duchesne, S. (2016). Exploring engagement in tasks in the language classroom. *Annual Review of Applied Linguistics, 36*, 50–72. https://doi.org/10.1017/S0267190515000094

Prensky, M. (2001). Digital natives, digital immigrants. *On the Horizon, 9*(5), 1–6. https://doi.org/10.1108/10748120110424816

Rowsell, J., & Walsh, M. (2011). Rethinking literacy education in new times: Multimodality, multiliteracies, & new literacies. *Brock Education Journal, 21*(1), 53–62. https://doi.org/10.26522/brocked.v21i1.236

Serafini, F. (2012). Expanding perspectives for comprehending visual images in multimodal texts. *Journal of Adolescent & Adult Literacy, 54*(5), 342–350. https://doi.org/10.1598/JAAL.54.5.4

Walsh, M. (2010). Multimodal literacy: What does it mean for classroom practice? *Australian Journal of Language and Literacy, 33*(3), 211–239.

Xu, J., Wang, Q., & Yang, L. (2023). Deep learning for multimodal educational data mining: Opportunities and challenges. *Journal of Educational Computing Research, 61*(2), 435–458. https://doi.org/10.1177/07356331221084722