



How Ready Are English Instructors for Multimodal Instruction in Vocational Colleges?

Muhammad Isa^{1*}, Elisa Ratih²

¹ English Study Program (D3), State Polytechnic of Madiun, East Java 61312, Indonesia

² Heavy Equipment Department (D3), Akademi Teknik Alat Berat Indonesia, East Java, 65124, Indonesia

Email: muhammadisa@pnm.ac.id, elisa@akabi.ac.id

ARTICLE INFO	ABSTRACT
Received: 04 Aug 2025	In response to the growing importance of English for Specific Purposes (ESP) in vocational higher education, multimodal teaching strategies have emerged as essential approaches to address diverse learning needs in the 21 st century. This study explores the preparedness of English instructors in Indonesian vocational colleges to integrate multimodal strategies into their instructional practices. Using a convergent mixed-method research design, the study collected data from ESP instructors across various majors via a semi-structured online questionnaire. The findings reveal that while most instructors demonstrate a positive disposition toward multimodal teaching and recognize its value in enhancing learner engagement, motivation, and 21 st -century skills, significant challenges persist. These include limited access to digital tools, time constraints in lesson preparation, uneven student digital literacy, and difficulty in assessing multimodal outputs. Moreover, despite alignment with the Merdeka Curriculum, some instructors remain uncertain about the curriculum's articulation of multimodal literacy. The study also highlights a preparedness gap stemming from inadequate training and institutional support. Nevertheless, instructors reported using a range of multimodal techniques, such as interactive media, group discussions, and digital projects, to foster communicative and task-based learning. The results underscore the necessity for structured professional development, institutional investment in digital infrastructure, and context-sensitive pedagogical models to enable sustainable multimodal integration. This research contributes to the growing body of literature on multimodal pedagogy by offering insights into instructors' perceptions, challenges, and the enabling conditions needed to promote effective multimodal English teaching in vocational education settings.
Accepted: 17 Oct 2025	
Published: 19 Dec 2025	
Keywords: Digital literacy, English for specific Purposes (ESP), Multimodal pedagogy, Teacher preparedness, Vocational education	
Doi: https://doi.org/10.59011/austronesian.4.3.2025.237-253	

*Corresponding Author

E-mail: muhammadisa@pnm.ac.id

Copyright © 2025 The Author(s)

CC BY-NC-SA 4.0 License

1. Introduction

In Indonesia, the rise of vocational higher education programs emphasizes the development of industry-specific skills, making English a vital component not only for general communication but also for enhancing technical competencies aligned with global labor-market needs. This instructional demand is often referred to as English for Vocational Purposes (EVP), which differs from General English in its emphasis on discipline-specific vocabulary and task-based communication tailored to workplace contexts (Widodo, 2016).

The distinctive nature of EVP requires language instruction that is both efficient and relevant to the learners' vocational fields, such as engineering, hospitality, finance, or information technology. As a result, the adoption of effective, engaging, and flexible teaching strategies becomes essential. One promising approach that aligns with 21st-century learning principles is multimodal pedagogy. Multimodal strategies integrate various semiotic resources, such as visual, auditory, spatial, gestural, and linguistic, to accommodate diverse learning preferences and to foster deeper engagement. This method allows students to interact with content through simulations, role plays, visual aids, and digital platforms, thus enabling meaningful learning experiences and the development of real-world language competencies (Liu, 2024).

Liao and Wu (2022) emphasizes that multimodal instruction not only enhances engagement but also supports collaborative learning environments by utilizing digital platforms for lesson planning, content sharing, and student-teacher interactions. Technological progress has made both synchronous and asynchronous communication more accessible, allowing teaching and learning activities to extend beyond the physical classroom. These affordances promote learner autonomy and peer interaction, enhancing the overall effectiveness of instruction. Several prior studies have confirmed the advantages of multimodal teaching in English language education. Choi and Yi (2016) found that multimodal approaches help learners grasp complex concepts and increase their confidence in using English. Similarly, Shi (2023) revealed that multimodal teaching has a positive psychological impact on students, with significant improvements in academic performance compared to traditional methods. Magnusson and Godhe (2019) further concluded that such approaches enhance student motivation, autonomy, and productive language skills. These findings support the growing interest in integrating multimodality into language teaching, especially in dynamic, skills-oriented learning environments such as vocational education.

Other studies have explored multimodal integration in specific contexts. For example, Sidabutar et al., (2021) and Aisyah and Nuraeni (2022) investigated textbook-based multimodal teaching, while Retno Wulan et al., (2022) examined the implementation of digital multimodal composition (DMC) in vocational schools, reporting increased student creativity and learning opportunities. Al-Farabi and Rohmah (2023) explored multimodal advertising texts as teaching media and found that they significantly improved students' communication skills. In the same vein, Ling and Mohammed (2024) reported that multimodal instruction improves listening skills and

enhances motivation in higher education. However, despite the promising outcomes of these multimodal initiatives, there remains a critical gap in the literature regarding the readiness and preparedness of English instructors, particularly in vocational higher education, to adopt and implement these strategies effectively. Many instructors continue to rely on monomodal, text-based approaches due to limited professional development, insufficient institutional support, and inadequate digital infrastructure (Guichon & Hauck, 2011). Moreover, some teacher training programs lack sufficient focus on multimodal teaching, leaving instructors unprepared for diverse and flexible classroom needs.

Furthermore, previous studies have explored multicultural critical thinking instruction (Darwin & Prasodjo, 2025), the implementation of the Merdeka Curriculum in secondary English classrooms (Ramli et al., 2025), and the effectiveness of monomodal reading strategies (Syaffi, 2021). None has systematically examined the readiness of vocational English instructors to integrate multimodal approaches aligned with the Merdeka Curriculum. This study seeks to fill this gap by focusing on instructors' preparedness, perceptions, and practical challenges in using multimodal tasks for English for Vocational Purposes (EVP) in Indonesian higher education.

This situation highlights several urgent questions: Are vocational English instructors equipped with the pedagogical knowledge and technical skills to design and implement multimodal instruction? Do institutional and curricular frameworks support the integration of varied meaning-making resources? While some educators express enthusiasm for multimodal learning, many feel constrained by systemic demands, including standardized assessments that prioritize traditional literacies over non-traditional, multimodal forms of communication.

Additionally, there is a growing disconnect between students' everyday exposure to multimodal texts in digital environments and the predominantly monomodal instruction they receive in formal education settings. This gap underscores the need to realign classroom practices with students' real-world literacies and to recognize multimodal teaching as a means to bridge this divide. In light of these challenges and opportunities, this study seeks to explore the preparedness of English instructors in vocational colleges to adopt multimodal strategies. By examining their perceptions, experiences, and institutional contexts, the research aims to identify barriers to implementation and to highlight the support systems required to effectively integrate multimodality into English for Vocational Purposes. Bridging this gap ensures vocational graduates are skilled and effective in multimodal communication.

Instructors' preparedness to adopt multimodal strategies is a multifaceted construct encompassing knowledge, attitudes, access to resources, and institutional support. In the context of English language teaching, Walsh (2011) emphasized that integrating multimodal strategies is not only about using media but also about enabling learners to understand how meaning is made across different modes. He emphasizes the critical role of multimodal literacy in classroom practice, particularly in English language instruction. He argues that effective teaching in the 21st century must go beyond traditional text-based approaches and integrate a variety of semiotic modes, including visual, audio, gestural,

and spatial elements. According to Walsh (2011), multimodality is not merely the use of technology or media in the classroom; rather, it involves a deliberate, pedagogically informed process in which teachers design learning experiences that align with how students interpret and create meaning in diverse ways. This perspective recognizes that different modes interact to support student engagement, deepen understanding, and cater to varied learning preferences.

Furthermore, Walsh highlights that teachers must not only be aware of how multimodal resources function but also be equipped to assess students' multimodal outputs effectively. He acknowledges the complexity of evaluating learning through multimedia formats and stresses the need for new forms of assessment that account for multiple modes of meaning-making. In vocational education, where learners must communicate both technically and practically, Walsh's framework provides a strong foundation for integrating multimodal strategies that mirror real-world communication tasks. His work thus underscores the importance of professional development and curriculum design that empower instructors to apply multimodal pedagogies in a purposeful and impactful way. However, many instructors still rely heavily on monomodal, text-based instruction due to limited training and institutional inertia (Fjørtoft, 2020; Godhe & Magnusson, 2017). Furthermore, Guichon and Hauck (2011) highlighted the psychological readiness of instructors, suggesting that even when resources are available, affective factors such as confidence, attitudes toward change, and prior teaching experience shape the adoption of multimodal methods. In the context of language education, monomodal instruction refers to teaching practices that primarily rely on a single mode of communication, typically text-based and linear, such as lectures, printed materials, or traditional writing tasks. This mode emphasizes reading and writing as isolated skills, often neglecting other semiotic resources.

Having considered the above elaboration, the primary objective of this study is to systematically examine the readiness and preparedness of English instructors in Indonesian vocational higher education to adopt and implement multimodal instructional strategies. Specifically, the research seeks to explore instructors' pedagogical knowledge, technical skills, and perceptions of integrating varied semiotic resources within the English for Vocational Purposes (EVP) framework. Furthermore, the study aims to identify the systemic barriers—including limited professional development, institutional inertia, and inadequate digital infrastructure—that hinder the shift from traditional monomodal instruction to flexible, 21st-century multimodal tasks. By addressing these factors, the study intends to bridge the gap between formal classroom practices and the real-world multimodal literacies required by the global labor market, ensuring that vocational graduates are effectively equipped for workplace communication.

2. Literature Review

The conceptual foundation of multimodal pedagogy is deeply rooted in social semiotics, which views meaning-making as the dynamic orchestration of multiple semiotic resources. Bezemer and Jewitt (2021) argue that multimodal learning involves

the orchestration of visual, spatial, gestural, and linguistic resources in learning environments that mirror authentic contemporary communication. This view is echoed by Serafini (2015), who emphasizes that educators must move beyond traditional print literacies to design explicit pedagogies that teach students how to interpret and produce multimodal texts. Further extending this, Lim (2021) highlights that embodied teaching integrates gestures and physical interaction as essential modes of communication, expanding the potential of multimodal instruction to engage learners cognitively and emotionally. Kalantzis and Cope (2020) reinforce this perspective by arguing that digital multimodal texts challenge teachers to cultivate students' critical literacies across various modes, demanding a deliberate pedagogical shift. Rohi and Nurhayati (2024) found that integrating various modes, such as linguistic, auditory, visual, and kinesthetic, can enhance EFL students' comprehension, retention, and motivation, demonstrating the practical benefits of multimodal strategies in Indonesian classrooms.

In line with this, Hidayah et al., (2025) demonstrate that implementing multiliteracies pedagogy in technology-supported speaking classes significantly develops students' multimodal competence and digital communication skills, aligning with the demands of the 21st-century workforce. Collectively, these insights highlight both the theoretical framework and the empirical evidence supporting the adoption of multimodal literacy, particularly in English for Vocational Purposes (EVP), where teacher preparedness and practical guidance remain key challenges. This distinction is crucial for vocational English contexts, where workplace communication often demands not only written proficiency but also skills in interpreting diagrams, using software, presenting visually, and interacting through digital media. Hence, multimodal pedagogy extends the scope of language teaching beyond the constraints of traditional monomodal approaches.

In vocational education, where learning outcomes are closely tied to industry relevance, multimodal approaches bridge abstract language concepts with practical, task-based learning. However, as noted by Hafner (2015), the successful implementation of such strategies hinges on instructors' ability to contextualize multimodal input within domain-specific language needs.

A systematic review by Rowsell and Walsh (2011) identified five key enablers of multimodal implementation: (a) curriculum alignment, (b) access to digital tools, (c) ongoing professional development, (d) collaborative planning among teachers, and (e) supportive leadership. These interconnected elements form the foundation for a sustainable, pedagogically meaningful multimodal learning environment. Curriculum alignment ensures that multimodal practices are not treated as add-ons but are integrated into learning outcomes and assessment strategies. Meanwhile, access to digital tools enables both teachers and students to engage with various media and modes, while continuous professional development equips educators with the knowledge and confidence to experiment with and adapt multimodal approaches. Regarding curriculum alignment and access to digital tools, Hiller and Zarnow (2024) demonstrated how digital platforms and communication technologies can facilitate the integration of theory and

practice in vocational training. Their study showed that providing access to multimodal tools enhances both instructional delivery and student engagement.

3. Method

This study employed a convergent mixed-methods design (Creswell & Plano Clark, 2018) to examine English instructors' preparedness to integrate multimodal strategies within Indonesian vocational English for Specific Purposes (ESP) contexts. This approach was selected to provide a more nuanced understanding of readiness by triangulating quantitative trends with rich, qualitative insights. While the quantitative component assesses general patterns of beliefs and institutional challenges, the qualitative data explores the "how" and "why" behind participants' experiences and pedagogical choices (Ivankova & Greer, 2015).

The participants (N = 16) were ESP instructors teaching across various vocational majors, including heavy equipment engineering, hospitality, business, and information technology, at vocational higher education institutions in Indonesia. A purposive sampling technique was used, specifically targeting instructors with experience in developing digital or multimodal materials or those currently navigating the transition to content-based instruction in vocational settings. This selection ensures that the data reflects the lived realities of practitioners at the intersection of language teaching and vocational demands.

Data were collected via a semi-structured online questionnaire (Appendix A). The instrument comprised 27 items categorized into four core dimensions:

1. Integration & Beliefs: Multimodal literacy within 21st-century learning.
2. Curriculum Alignment: Perceptions of multimodality under the *Kurikulum Merdeka*.
3. Professional Background: Prior experience and professional development.
4. Ecological Challenges: Institutional and classroom-level barriers.

The quantitative strand consisted of 22 closed-ended statements measured on a five-point Likert scale (1 = Strongly Disagree; 5 = Strongly Agree). The qualitative strand comprised five open-ended questions (Items 17, 24, 25, 26, and 27) designed to elicit detailed descriptions of teaching strategies, media usage, and specific implementation hurdles.

The instrument was grounded in Walsh's (2011) framework of multimodal literacy, which emphasizes the interplay of five semiotic modes: linguistic, visual, auditory, gestural, and spatial. To ensure content validity, the questionnaire was reviewed by an expert in multimodal pedagogy. Participation was entirely voluntary, and confidentiality was maintained through data anonymization.

Quantitative data were analyzed using descriptive statistics (means and percentages) to identify global trends across the four dimensions. Qualitative responses were analyzed using Thematic Analysis (Braun & Clarke, 2006), a process that involves coding and the construction of categories. Finally, the results of both strands were integrated during the discussion phase to provide a holistic assessment of multimodal readiness. Data from 16

respondents were analyzed to determine the level of multimodal readiness among instructors. Scores were calculated on a 5-point Likert scale, and percentages were derived from the formula $(\text{mean}/5 \times 100)$. For Item 15, which originally used a frequency scale (Sometimes, Often, Always), the responses were converted into numerical scores (1–5) for consistency.

The questionnaire was designed to examine four core dimensions: (1) Integration of multimodal literacy within the framework of 21st-century learning; (2) Perceptions of multimodal literacy under the Merdeka Curriculum; (3) Professional development and prior experience in implementing multimodal teaching strategies; and (4) Institutional and classroom-level challenges in using multimodal approaches. The questionnaire was grounded in prominent theoretical frameworks of multimodal literacy, drawing on Walsh (2011). These frameworks emphasize the importance of recognizing multiple semiotic modes, visual, auditory, gestural, spatial, and linguistic, in constructing meaning in educational settings. To ensure content validity, the instrument was reviewed by an academic expert in multimodal pedagogy and revised accordingly to align with the study objectives and participant context.

The detailed categorization of the instrument items is presented in Table 1. Responses to the open-ended questions were analyzed using thematic analysis, involving coding and categorizing the data to identify emerging patterns relevant to the research objectives.

Table 1. The Research Instrument Grid

No.	Descriptions	Item Number
1.	Integration and Beliefs about Multimodality in 21 st Century Learning	1,2,3,4,5
2.	Teachers' Views on Multimodal Literacy According to Merdeka Curriculum	6,7,8,9,10
3.	Teachers' Experience and Professional Development Related to Multimodal Literacy	11,12,13,14,15,16
4.	Challenges Encountered by Teachers in the Integration of Multimodal Literacy	18,19,20,21,22,23

Overall, the chosen methodological approach facilitated a comprehensive exploration of instructor readiness for multimodal English teaching in vocational contexts, enabling the study to capture both the breadth and depth of the phenomenon under investigation.

4. Results and Discussion

This study explored the preparedness and perceptions of English instructors in vocational colleges regarding the use of multimodal strategies in their teaching practices. A total of 27 questionnaire items, both closed- and open-ended, were distributed to gather insights into instructors' implementation, beliefs, challenges, and needs regarding multimodal literacy. Data from 16 respondents were analyzed to determine the level of multimodal readiness among instructors. Scores were calculated on a 5-point Likert scale, and percentages were derived from the formula $(\text{mean}/5 \times 100)$. For Item 15, which originally used a frequency scale (Sometimes, Often, Always), the responses were

converted into numerical scores (1–5) for consistency.

4.1. Summary of Dimensions

The quantitative analysis categorized the instructors' readiness into four primary dimensions, as summarized in Table 2.

Table 2. Summary of Readiness Scores per Dimension

Dimension	Items	Global Mean	Percentage	Category
Multimodal Integration & Beliefs	1-5	4.29	85.75%	Very High
Perceptions of Kurikulum Merdeka	6-10	3.75	75.00%	High
Experience & Professional Development	11-17	3.51	70.21%	High
Challenges in Integration	18-23	3.44	68.75%	High

4.2. Quantitative Item Analysis

Table 2 presents a detailed breakdown of the mean scores and categories for each quantitative item.

Table 3. Summary of Quantitative Items

Item	Mean	Percentage	Category	Item	Mean	Percentage	Category
1	4.31	86.25%	Very High	12	3.12	62.50%	High
2	4.31	86.25%	Very High	13	3.56	71.25%	High
3	4.62	92.50%	Very High	14	4.19	83.75%	Very High
4	4.19	83.75%	Very High	15	3.50	70.00%	High
5	4.00	80.00%	High	16	4.06	81.25%	Very High
6	3.94	78.75%	High	18	3.38	67.50%	High
7	3.69	73.75%	High	19	3.25	65.00%	High
8	3.88	77.50%	High	20	3.88	77.50%	High
9	3.56	71.25%	High	21	3.38	67.50%	High
10	3.69	73.75%	High	22	3.81	76.25%	High
11	2.62	52.50%	Moderate	23	2.94	58.75%	Moderate

4.3. Narrative Summary of Quantitative Results

- *Multimodal Integration & Beliefs*: This dimension achieved the highest score (Mean = 4.29; 85.75%), indicating that instructors hold strong convictions regarding the benefits of multimodal learning. Notably, Item 3 recorded the highest score in the entire survey (Mean = 4.62; 92.50%), suggesting a consensus that students benefit significantly from learning through diverse communicative modes.
- *Perceptions of Kurikulum Merdeka*: Categorized as "High" (Mean = 3.75; 75.00%), instructors generally view the *Kurikulum Merdeka* as a framework that supports multimodal literacy integration. However, Item 9 received a lower score (Mean = 3.56; 71.25%), highlighting a need for more robust access to implementation guidelines.

- *Experience & Professional Development*: This dimension scored "High" (Mean = 3.51; 70.21%). Instructors are actively utilizing digital platforms and seeking new strategies. While Item 14 (Mean = 4.19) showed strong proficiency in tools such as Canva, Padlet, and video editors, Item 11 yielded a lower score (Mean = 2.62; 52.50%), indicating a lack of formal training specifically focused on multimodal literacy.
- *Challenges in Integration*: With a mean of 3.44 (68.75%), implementation challenges remain significant. The most prominent obstacles include variations in students' digital literacy (Item 20, Mean = 3.88) and the need for institutional leadership support (Item 22, Mean = 3.81).

4. Thematic Analysis of Qualitative Data

- *Professional Training (Item 17)*: Half of the respondents reported having no specialized training in multimodal literacy. While some had participated in workshops on digital media, gamification, or student-centered learning, the experiences were inconsistent and often lacked a specific focus on multimodal pedagogy.
- *Key Challenges (Item 24)*: The primary challenge, cited by 8 respondents, pertains to student readiness, including variations in digital literacy and language proficiency. Other significant barriers include limited resources and facilities (7 respondents) and time constraints for preparation (4 respondents).
- *Instructional Strategies (Item 25)*: Common strategies include collaborative work, peer teaching, and Jigsaw discussions. Instructors also frequently employ visual-based strategies (infographics, games, digital media), reflecting an adoption of active learning, though the depth of multimodal design varies.
- *Media Utilization (Item 26)*: Visual media (posters, diagrams, charts) and audio-visual media (YouTube, TED Talks, interactive videos) are widely used. Audio and oral modes, such as podcasts and speech recordings, also feature strongly in the classroom.
- *Classroom Activities (Item 27)*: Presentations and group discussions are the most frequent activities. There is a growing trend toward multimodal product-based projects (e.g., creating digital content or posters). However, there is a clear need for standardized assessment rubrics and formal pedagogical guidance to strengthen these practices.

5. Discussion

The findings of this study indicate that the majority of English instructors in vocational colleges hold a positive attitude toward the implementation of multimodal teaching strategies. This result aligns with the study by Yuniarti & Yulian (2022), who emphasized that integrating multimodal approaches can enhance student engagement and learning outcomes in English language classrooms. Araújo further highlighted that multimodal literacy, which encompasses the ability to interpret and construct meaning

through various modes of communication such as visual, auditory, and digital technologies, can enrich students' learning experiences. Furthermore, the results of this study align with and extend previous findings. While Choi & Yi (2016); Magnusson & Godhe (2019) This study emphasizes the benefits of multimodal strategies for learners' confidence and engagement. It adds the perspective of teacher receptiveness, especially in the context of vocational education. Similar to Guichon & Hauck (2011) A disparity in teacher confidence emerged, revealing that despite general enthusiasm, institutional support and training remain inadequate. Thus, the findings of this study indicate that while English instructors in Indonesian vocational colleges show an encouraging readiness to integrate multimodal strategies, there are practical and theoretical tensions that must be addressed. The positive disposition aligns with global trends in 21st-century education (Yuniarti & Yulian, 2022), but the local context reveals unique challenges that widen the implementation gap.

Despite the growing awareness of the importance of multimodal literacy, several instructors in this study expressed uncertainty regarding the conceptualization of multimodality as articulated in the *Merdeka Curriculum*. A significant insight from the findings is the tension between *teacher autonomy* and *policy clarity*. Although the *Merdeka Curriculum* affords pedagogical flexibility, the lack of explicit guidelines for designing, implementing, and assessing multimodal tasks leaves many instructors unsure. As one participant expressed, "*I want to try more digital tasks but often hesitate because there is no clear standard to follow for scoring and alignment with the syllabus.*" This indicates that while teachers have the autonomy to innovate, policy ambiguity can restrict their confidence and consistent application. This observation echoes the findings of Yunus et al. (2022), who found that although teachers may possess a basic understanding of multimodal literacy, there remains a significant need for targeted professional development and training to deepen their conceptual knowledge and facilitate effective implementation in instructional practice. The empirical gap remains pronounced: "No prior study in Indonesia has systematically explored the alignment between the *Merdeka Curriculum* and multimodal ESP instruction in vocational colleges." This underscores the need for curriculum developers to develop clearer frameworks that connect multimodal literacies to task-based vocational learning outcomes. Without explicit scaffolding, teachers' efforts risk remaining fragmented.

Moreover, the challenges faced by instructors in this study, such as time constraints in planning multimodal lessons, variations in students' levels of digital literacy, and the complexity of assessing multimodal texts, are consistent with prior literature. For instance, Ross et al., (2020) argued that the assessment practices are often insufficient to capture how multimodal works represent disciplinary knowledge in complex ways that make technical and compositional assessment criteria do not always address the richness and complexity of multimodal work. In terms of evaluation, this study confirms that assessing multimodal student work presents a distinct challenge compared to monomodal, text-based assessment. Traditional written tasks are judged by familiar grammar, coherence, and vocabulary rubrics. In contrast, multimodal projects require evaluation of

visual design, audio integration, creativity, and technological skills, for which standardized rubrics are scarce. As one participant highlighted, *“It’s confusing to judge how much weight should go to language accuracy versus creativity when students make videos or digital posters.”* This points to a vital theoretical implication: assessment practices must evolve to recognize new literacies rather than merely adapt old ones.

Despite these challenges, the instructors demonstrated a strong commitment to integrating multimodal teaching strategies, including group discussions, videos, presentations, and creative digital projects. This pedagogical orientation is supported by the findings of Yunus et al. (2022), which suggests that the use of diverse media and digital tools in English language instruction can foster greater student motivation and engagement, while also promoting critical 21st-century skills such as collaboration and critical thinking. Nevertheless, consistent with Hafner (2015) In the work, instructors demonstrated creative agency by incorporating multimodal activities such as digital storytelling and collaborative presentations, highlighting their adaptability amid systemic constraints. Overall, the results of this study underscore the importance of multimodal literacy in vocational English education and highlight the need for sustained support through professional development, adequate resources, and responsive educational policies to enable instructors to implement multimodal strategies effectively.

The findings of this study resonate with the broader global discourse on the urgent need to enhance teacher capacity in digital and multimodal pedagogy, particularly in light of the technological shifts and evolving literacy practices of the 21st century. In Indonesia, where vocational institutions are under increasing pressure to produce graduates with practical, industry-relevant skills, multimodal literacy is no longer an optional enhancement but a core competency. Despite progressive policy frameworks such as the Merdeka Curriculum, institutions' readiness to implement these pedagogical innovations often falls short, hindered by gaps in infrastructure, training, and systemic support. To bridge this divide, it is imperative to establish structured and ongoing professional development programs that address specific pedagogical needs, ranging from the design of multimodal instructional materials aligned with English for Specific Purposes (ESP) objectives to the application of appropriate assessment frameworks for evaluating multimodal outputs. It aligns with Wahyudi & Jufrizal (2023), the study showed that regarding course preferences, 83% of the respondents expressed a desire to study English, which aligns with the specific materials required in the engineering field. Additionally, 92% indicated a need to develop strong negotiation and instruction comprehension skills to support their future professional roles. Moreover, institutions must invest in building accessible digital content repositories and fostering communities of practice where instructors can engage in peer mentoring and reflective teaching. Importantly, there remains a critical need for further research to develop localized, context-sensitive models of multimodal instruction that are attuned to the unique cultural, infrastructural, and pedagogical conditions of vocational education in Indonesia.

Although this study provides valuable insights into the preparedness of English instructors in vocational colleges to adopt multimodal teaching strategies, several

limitations must be acknowledged. First, the study relied on self-reported data collected through online questionnaires, which may introduce social desirability bias or limit the depth of responses. While open-ended questions were included, they may not capture the full complexity of instructors' experiences as effectively as in-depth interviews or classroom observations. Second, the sample was limited to vocational colleges in Indonesia and may not fully represent the diversity of institutions or teaching contexts across other regions or countries. This affects the generalizability of the findings. Third, the study did not directly measure classroom implementation or student learning outcomes, which are critical in evaluating the actual effectiveness of multimodal strategies. Finally, the research focused primarily on instructors' perspectives; future studies could benefit from incorporating the voices of students and administrators to obtain a more holistic understanding of institutional readiness for multimodal instruction.

6. Conclusion

This study provides valuable insights into the preparedness and perceptions of English instructors in vocational colleges regarding the integration of multimodal strategies in language teaching. The findings demonstrate a generally positive orientation among instructors toward multimodal pedagogy, with many reporting frequent use of diverse semiotic resources, such as visual, auditory, gestural, spatial, and linguistic in their instructional practices. Such integration is perceived to be instrumental in enhancing student engagement, fostering 21st-century skills, and aligning English instruction with the demands of contemporary literacy.

Moreover, the *Merdeka* Curriculum was widely regarded as supportive of multimodal literacy, offering pedagogical flexibility that enables instructors to tailor instructional approaches to learners' needs. Nevertheless, a small portion of instructors indicated limited awareness of the curriculum's conceptualization of multimodality, suggesting a need for greater curricular clarity and professional development. Despite these positive dispositions, instructors encountered several challenges that hinder the full realization of multimodal teaching. These include time constraints, disparities in students' digital competencies, difficulties in assessing multimodal outputs, and limited institutional support and resources. Nonetheless, instructors demonstrated resilience and creativity by adopting a variety of multimodal strategies, leveraging both traditional and digital tools to foster communicative and collaborative learning environments.

Furthermore, this study offers important empirical insights and highlights a critical gap: the alignment between the flexible *Merdeka* Curriculum and the practical realities of multimodal ESP instruction has not been systematically studied before in Indonesia. Addressing this gap is vital to transform the promise of 21st-century learning into a sustainable classroom practice. Overall, the research calls for an urgent recalibration of policy, professional development, and curriculum design to ensure that multimodal literacy is not an add-on but a fully integrated feature of vocational English education. Clearer policy articulation, teacher training, and robust assessment frameworks are essential steps toward ensuring Indonesian vocational graduates are equipped with

relevant multimodal communication skills for today's dynamic work environments.

Taken together, the findings underscore the growing importance of multimodal literacy in vocational English education. To ensure effective and equitable implementation, sustained institutional support is essential—particularly through targeted training, accessible technological infrastructure, and clear policy guidance. Future research may further explore longitudinal impacts of multimodal instruction on student outcomes and investigate scalable models of professional development tailored to the vocational context.

Conflict of interest

The authors hereby declare that there are no conflicts of interest that could have appeared to influence the work reported in this paper.

Authors' contribution

The authors state that the author made substantial contributions to the conception and design of the study. The authors were responsible for data analysis, interpretation, and discussion of the results. The authors read and approved the final manuscript.

References

- Aisyah, S., & Nuraeni, N. (2022). The Use of a Digital Multimodal Textbook to Foster Students' English Proficiency. *Jo-ELT (Journal of English Language Teaching) Fakultas Pendidikan Bahasa & Seni Prodi Pendidikan Bahasa Inggris IKIP*, 9(2), 160. <https://doi.org/10.33394/jo-elt.v9i2.6234>
- Braun, V., & Clarke, V. (2006). Using Thematic Analysis in Psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp0630a>.
- Bezemer, J., & Jewitt, C. (2021). *Introducing Multimodality* (2nd ed.). Routledge.
- Choi, J., & Yi, Y. (2016). Teachers' Integration of Multimodality Into Classroom Practices for English Language Learners. *TESOL Journal*, 7(2), 304–327. <https://doi.org/10.1002/tesj.204>.
- Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and Conducting Mixed Methods Research* (3rd ed.). SAGE Publications.
- Darwin, & Prasadjo, P. (2025). Critical Thinking Skills Program: EFL Classroom Multicultural Planning and Teaching Practice in Higher Education. *Celtic: A Journal of Culture, English Language Teaching, Literature and Linguistics*, 12(1), 88–102. <https://doi.org/10.22219/celtic.v12i1.36499>
- Fjørtoft, H. (2020). Multimodal Digital Classroom Assessments. *Computers and Education*, 152. <https://doi.org/10.1016/j.compedu.2020.103892>
- Godhe, A., & Magnusson, P. (2017). Multimodality in Language Education – Exploring the Boundaries of Digital Texts. *International Conference on Computers in Education*, 845–854.
- Guichon, N., & Hauck, M. (2011). Editorial: Teacher Education Research in CALL and CMC: More in Demand than Ever. *ReCALL*, 23(3), 187–199. <https://doi.org/10.1017/S0958344011000139>

- Hafner, C. A. (2015). Remix Culture and English Language Teaching: The Expression of Learner Voice in Digital Multimodal Compositions. *TESOL Quarterly*, 49(3), 486–509. <https://doi.org/10.1002/tesq.238>
- Hidayah, J., Prihantoro, Apriani, E., Polisda, Y., Arnanda, F. S., & Oktavia, N. (2025). Facilitating Students' Multimodal Competence Development in Technology- Supported Speaking Class through Multiliteracies Pedagogy. *Journal of English Education and Teaching (JEET)*, 9(1), 172–189. <https://doi.org/10.33369/jeet.9.1.172-189>.
- Hiller, F., & Zarnow, S. (2024). Fostering Innovative Learning and Satisfaction in Virtual Teamwork: Shedding Light on Apprentices. *International Journal for Research in Vocational Education and Training*, 11(2), 171–199. <https://doi.org/10.13152/IJRJET.11.2.2>
- Ivankova, N. V., & Greer, K. (2015). Mixed Methods Research in Language Teaching and Learning. *The Routledge Handbook of Research Methods in Applied Linguistics*. Routledge
- Kalantzis, M., & Cope, B. (2020). *Multiliteracies for a Digital Age* (2nd ed.). Routledge. Language Learners. *TESOL Journal*, 7(2), 304–327. <https://doi.org/10.1002/tesj.204>
- Liao, C. H., & Wu, J. Y. (2022). Deploying Multimodal Learning Analysis Models to Explore the Impact of Digital Distraction and Peer Learning on Student Performance. *Computers and Education*, 190. <https://doi.org/10.1016/j.compedu.2022.104599>
- Lim, F. V. (2021). *Designing Learning Materials for Multimodal Literacy*. Routledge.
- Ling, Z., & Mohammed, L. A. (2024). A Study on the Effect of Multimodal Teaching on Higher Vocational English Listening. *Curriculum and Teaching Methodology*, 7(2), 181–186. <https://doi.org/10.23977/curtm.2024.070227>
- Liu, Y. (2024). Research on Multimodal Method in English Teaching: Theory and Practice. *Curriculum and Teaching Methodology*, 7(2), 115–121. <https://doi.org/10.23977/curtm.2024.070218>
- Magnusson, P., & Godhe, A. L. (2019). Multimodality in Language Education – Implications for Teaching. *Designs for Learning*, 11(1), 127–137. <https://doi.org/10.16993/dfl.127>
- Al-Farabi, M. R. F., & Rohmah, Z. (2023). Teaching Multimodal Literacy Through Integrated Learning of Advertisement Texts. In *FKIP Universitas Lambung Mangkurat Banjarmasin*, 6(1), 1-10. <https://dx.doi.org/10.20527/jetall.v6i1.15340>
- Ramli, Jawan, J. N., Ridwan, R., Kusmaryani, W., & Rosmayanti, V. (2025). English Language Teaching on Merdeka Curriculum at the Border Area of Indonesia Malaysia. *Celtic: A Journal of Culture, English Language Teaching, Literature and Linguistics*, 12(1), 213–229. <https://doi.org/10.22219/celtic.v12i1.40019>
- 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>
- Ross, J., Curwood, J. S., & Bell, A. (2020). A Multimodal Assessment Framework for Higher Education. *E-Learning and Digital Media*, 17(4), 290–306. <https://doi.org/10.1177/2042753020927201>
- Rowell, 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. (2015). *Reading the Visual: An Introduction to Teaching Multimodal Literacy*. Teachers College Press. 208-215
- Shi, L. (2023). Empowering Language Education: Integrating Translanguaging, Multimodality,

- and Dialogic Teaching in Chinese Private Universities. *International Journal of Sociologies and Anthropologies Science Reviews*, 3(6), 117–128. <https://doi.org/10.60027/ijssar.2023.3471>
- Sidabutar, U., Sinaga, N. T., Sitorus, N., & Lestari, F. D. (2021). Multimodal Analysis on a Vocational English Book and its Effect on Students' English Proficiency. *Linguistics and Culture Review*, 5(S3), 1651–1665. <https://doi.org/10.21744/lingcure.v5ns3.1951>
- Syafii, M. L. (2021). The Implementation of the Story Mapping Strategy to Enhance Students' Reading Comprehension. *Celtic: A Journal of Culture*, 8(1). <https://doi.org/10.22219/celtic.v8i1.16161>
- Wahyudi, W., & Jufrizal, J. (2023). The Need of English for Specific Purposes in Vocational High School. *AL-ISHLAH: Jurnal Pendidikan*, 15(3), 3084–3093. <https://doi.org/10.35445/alishlah.v15i3.3047>
- Walsh, M. (2011). *Multimodal literacy: Researching classroom practice*. Sydney, Australia: Primary English Teaching Association.
- Widodo, H. P. (2016). Teaching English for Specific Purposes (ESP): English for Vocational Purposes (EVP). In *English Language Education* (Vol. 5, pp. 277–291). Springer Science and Business Media B.V. https://doi.org/10.1007/978-3-319-38834-2_19
- Yuniarti, Y., Yulian, R., & Yuniarti, Y. (2022). Digital Story Telling Based on Multimodal Elements on EFL Learners' Speaking Performance. *JPI (Jurnal Pendidikan Indonesia)*, 11(2), 308–316. <https://doi.org/10.23887/jpiundiksha.v11i2.40217>
- Yunus, M., Suciati, Ritonga, M., & Kumar, T. (2022). Multimodal teaching practices for EFL teacher education: An Action-based Research Study. *Journal for Educators, Teachers and Trainers*, 13(1), 142–148. <https://doi.org/10.47750/jett.2022.13.01.016>

Appendix A: Research Instrument Multimodal Readiness Questionnaire for ESP Instructors

Part I: Demographic Information

- **Name (Optional/Anonymized):**
- **Institution:**
- **Vocational Major Taught:** (e.g., Hospitality, Engineering, IT, etc.)
- **Teaching Experience (Years):**

Part II: Quantitative Measures (Likert Scale)

Instructions: Please indicate your level of agreement with the following statements on a scale of 1 to 5.

(1 = Strongly Disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, 5 = Strongly Agree)

No	Item Statement	1	2	3	4	5
Dim 1	Integration & Multimodal Beliefs					
1	I believe multimodal strategies are essential for 21st-century ESP teaching.					
2	Using multiple modes (visual, audio, etc.) improves student engagement.					
3	Students benefit from learning through various communication modes.					
4	Multimodal literacy is as important as traditional linguistic literacy.					
5	I feel confident in my ability to design multimodal learning materials.					
Dim 2	Perceptions of Kurikulum Merdeka					
6	<i>Kurikulum Merdeka</i> provides flexibility for multimodal integration.					
7	The current curriculum supports the development of digital literacies.					
8	Multimodal literacy aligns with the "Merdeka Belajar" philosophy.					
9	I have clear access to guidelines for implementing multimodal pedagogy.					
10	The curriculum encourages the use of diverse semiotic resources.					
Dim 3	Experience & Professional Development					
11	I have received formal training specifically on multimodal literacy.					
12	I frequently participate in workshops related to digital teaching media.					
13	I actively seek new digital strategies for my ESP classes.					
14	I am proficient in using tools like Canva, Padlet, or video editors.					
15*	Frequency of using digital platforms in class (Converted to 1–5 scale).					
16	My previous teaching experience helps me adapt to multimodal modes.					
Dim 4	Institutional & Classroom Challenges					
18	Lack of high-speed internet hinders my multimodal implementation.					
19	Preparing multimodal materials is more time-consuming than traditional ones.					
20	Student digital literacy levels vary significantly in my classroom.					
21	There is a lack of high-quality, pre-made multimodal ESP resources.					
22	I receive strong support from institutional leadership for these methods.					
23	Large class sizes make multimodal activities difficult to manage.					

Part III: Qualitative Measures (Open-Ended Questions)

Instructions: Please provide brief descriptions based on your classroom experience.

- **Item 17:** Describe any specific training or workshops you have attended regarding multimodal literacy or digital pedagogy.
- **Item 24:** In your opinion, what are the primary challenges when implementing multimodal strategies in a vocational English context?
- **Item 25:** What specific instructional strategies (e.g., Jigsaw, Peer Teaching, Project-Based Learning) do you use to facilitate multimodal learning?
- **Item 26:** List the types of media (visual, audio, audio-visual) you most frequently utilize in your ESP teaching.
- **Item 27:** Describe a typical classroom activity where students are required to produce or interact with multimodal content.