



PROBING THE ALIGNMENT BETWEEN MULTIMODAL TEACHING PRACTICE AND STUDENTS' LEARNING PREFERENCES IN INDONESIAN EFL CLASSROOM

Novia Trisanti¹, Henrikus Joko Yulianto², Daud Jiwandono³, Almira Irwaniyanti Utami⁴

^{1,3}English Education Study Program, Faculty of Languages and Arts, Universitas Negeri Semarang, Indonesia

^{2,4}English Literature Study Program, Faculty of Languages and Arts, Universitas Negeri Semarang, Indonesia

Corresponding author's email: novia.santi@mail.unnes.ac.id

APA Citation: Trisanti, N., Yulianto, H. J., Jiwandono, D., & Utami, A. I (2026). Probing the alignment between multimodal teaching practice and students' learning preferences in Indonesian EFL classroom. *English Review: Journal of English Education*, 14(2), 589-600. <https://doi.org/10.25134/englishreview.v14i2.210>

Received: 17-01-2026

Accepted: 25-03-2026

Published: 09-06-2026

Abstract: In Indonesia, teaching and learning English are confronted with various obstacles in aligning instructional approaches with students' learning preferences and multimodal literacy expected by the curriculum. Multimodal pedagogy has been widely promoted in EFL contexts. However, few research have studied the alignment between teachers' multimodal teaching approaches and students' preferred learning modalities in authentic classrooms. This research explored the perceived alignment of multimodal teaching with student's learning preferences in the Indonesian secondary EFL classroom. A mixed-method approach was employed in this study with 20 EFL teachers and 60 junior high school students in Semarang Regency, Indonesia. Data were obtained by Likert scale questionnaire and open-ended response and analyzed by descriptive statistics and qualitative analysis. The results showed that students had the highest preference for kinesthetic learning ($M = 7.78$) and video-based media (70%), while the lowest preference was for audio-based learning ($M = 3.04$). Students had generally good perception of teachers' use of multimodal media. However, the convergence of instructional practices to students' chosen modalities was only partial as indicated by the moderate mean scores ranging from 3.45 to 3.58. The qualitative data findings to support the quantitative analysis results also reflect that the students preferred interactive and visually aided learning activities since they raised engagement and understanding. This study reveals multimodal pedagogy has been applied in Indonesian EFL classrooms, although further pedagogical adaptation and teacher support are needed to meet learner diversity and maximize multimodal instruction.

Keywords: *multimodal literacy; learning preferences; teaching practices; student preferences*

INTRODUCTION

In recent years, the field of English language teaching has been progressively emphasizing the integration of digital technology and multimodal literacy practices in response to the demands of 21st-century education (Humairoh et al., 2025; Liu & Lim, 2024; Pratiwi & Trisanti, 2025). The development of visual, aural, spatial and digital modalities of communication has underlined the significance of multimodal literacy, that is, the ability to perceive and produce meaning through different semiotic resources, such as images, videos, gestures and digital texts (Collet et al., 2026; Puspitasari et al., 2025). This trend is especially important for EFL contexts, such as in

Indonesia, where learners are increasingly exposed to multimodal information through social media, online platforms and digital learning environments. Thus, modern English language instructional process is aligned with student-centered and multimodal pedagogies that employ a range of resources such as films, interactive media, and visual materials to foster engagement and comprehension. Previous studies have indicated the positive effects of multimodal instruction on learners' motivation, understanding, and meaning-making processes in language learning (Guo et al., 2024; Tandiana et al., 2023; Zhu et al., 2026). Furthermore, the convergence of multimodal resources is in line

with communicative and context-based language teaching methodologies that promote authentic and meaningful learning experiences (Dwi Jayanti & Damayanti, 2023; Q. Lei & Zhang, 2024; Yi et al., 2017)

Having meaningful learning experiences in EFL classes using multimodal instructions, students may like different ways to learn, like hearing, seeing, reading and writing, or doing things (Collet et al., 2026; Dayanti et al., 2023; Q. Lei & Zhang, 2024; Sitokkonni, 2025). Knowing these sensitivities might help teachers create more interesting and easy-to-understand lessons that encourage student participation and independence. The connection between how teachers teach and how students like to learn has become a significant dilemma in language education. Many studies have linked multimodal teaching to better engagement and comprehension (Humairoh et al., 2025; Ulu & Cahit, 2025; W. Wang et al., 2026; Zhu et al., 2026), but most of those studies have only looked at learning outcomes and not how multimodal practices fit with what students want in the real classroom.

A lot of research has been done on digital literacy and integrating technology, but not much has been done on how teachers' use of multiple teaching methods matches up with how students like to learn in real classrooms (Indrayani et al., 2024; Nurhaliza & Sari, 2023; Zhu et al., 2026). Also, a lot of the studies that have been done before only look at things from the point of view of teachers or students. This means that we don't really understand how to balance classrooms from different points of view. In educational study, learning preferences are still being discussed (El-Saftawy et al., 2024; Norhasanah et al., 2022; Sumardi & Tahir, 2025), but new ideas say that they should be seen as adaptable and situational instead of fixed learner traits.

Alongside the progress of multimodal pedagogy, the concept of learner diversity, particularly concerning learning preferences, has received considerable attention. Learning preferences are categorized into modalities, such as visual, auditory, reading/writing, and kinesthetic, which demonstrate the diverse methods through which learners opt to acquire and interpret information (El-Saftawy et al., 2024; Herlita, 2025; Norhasanah et al., 2022; Nurhaliza & Sari, 2023). The importance of understanding learners' preferences have been emphasized in education system, including the in a curriculum, since this insight can aid educators in developing more successful and customized instructional

strategies (Nurhaliza & Sari, 2023; Saputra & Dewi, 2024; Sitokkonni, 2025). Therefore, teaching methods should be adaptable to accommodate the diverse traits of learners as becoming the importance of individual variability in language acquisition.

The concept of learning preferences remains a contentious issue since empirical data supporting the effectiveness of aligning teaching styles with learning styles is few, recommending prudence in the simplistic use of this notion (Amjad et al., 2024; Putri Nur Wandah et al., 2024; Wiangga & Febriyanto, 2024). Critical examination of the prevalent utilization of learning style models, emphasizing the theoretical contradictions and their ramifications for practical implementation have been asserted in learners' preferences related to the way they learn (Pillai & Khan, 2025). Despite these criticisms, the notion of learning preferences continues to be widely employed as a pedagogical concern, particularly when viewed as flexible and context-dependent rather than as fixed traits.

Another previous studies highlight the ongoing difficulties encountered in English language education in Indonesia, notably stressing the imperative to adapt teaching methods to meet the evolving needs of students and technological progress (Mark & Bernadeth, 2024; Pikhart et al., 2024; Wiangga & Febriyanto, 2024). Meanwhile, the efficacy of digital and multimodal methodologies, for instance particularly related to visual media, infographics, genre-based digital storytelling, etc., which can enhance learner engagement and promoting language acquisition (Carcamo & Pino, 2025; Hidayati et al., 2024; Lei & Zhang, 2024; Pratiwi & Trisanti, 2025; Tsai, 2025). Nonetheless, some current researches mostly focus on the implementation and outcomes of multimodal practices, rather than examining the degree to which these practices align with the particular learning preferences of students (de Oliveira, 2024; Derakhshan et al., 2026; Guo et al., 2024; Quijada & Tagle, 2025; Sakdiyah et al., 2026; Wang et al., 2023; Yi et al., 2017). Therefore, the importance of multimodal discourse analysis in clarifying the mechanisms by which meaning is generated across different modes; yet, its application in classroom-based research remains limited. In a similar vein, while theoretical frameworks for multimodal literacy are well-established, there is a requisite for further empirical research to investigate the implementation of these frameworks in real educational contexts (Guo et al., 2024; Pillai &

Khan, 2025; Pratiwi et al., 2024; Stewart, 2023; Tandiana et al., 2023). This underscores a divergence between theoretical progress and practical implementation in the educational environment.

Multimodal pedagogy has a great potential to promote language learning, but the extent of its success can be contingent on the degree to which instructional techniques align with learners' preferences and classroom demands. Few research explicitly investigated this alignment, especially in secondary EFL situations. Most of prior research has examined higher education settings, with minimal understanding of multimodal pedagogical approaches in secondary classrooms with learners of varying cognitive, emotional and developmental characteristics. (Humairoh et al., 2025; Lei & Zhang, 2024; Li & Dell-Jones, 2023; Stewart, 2023; Trisanti, et., al., 2024). In fact, classroom realities such as curriculum expectations, teacher readiness, and resource availability emphasize the need for more context-specific investigations. Thus, exploring the relationship between multimodal teaching and students' learning preferences should provide useful insights for improving EFL instructional design and promoting teacher professional development.

Therefore, this study explores students' learning preferences and their impressions of multimodal teaching approaches in EFL classes. The study specifically explores the extent to which students perceive alignment between teachers' multimodal teaching approaches and their preferred means of learning. This study integrated quantitative and qualitative data from teachers and students to offer empirical understandings of multimodal pedagogy and learner's preference in EFL classroom contexts.

METHOD

This study adopted a mixed-method methodology to look into the alignment of students' learning preferences and multimodal teaching practices. The mixed-method technique was utilized to collect quantitative and qualitative data on how students perceived multimodal teaching in connection to their preferred learning styles (Cresswell & Cresswell, 2023). Quantitative data were used to determine trends in students' preferences and perceptions. Qualitative responses gave more in-depth explanations of students' classroom experiences and perspectives of multimodal instruction. The study was conducted in five public junior high schools in

Semarang Regency of Central Java. It involved two distinct groups: a cohort of 20 EFL teachers and a sample of 60 junior high school students in Semarang regency. The sampling of students' participants were from the classes of collaborating teachers and had undergone multimodal learning activities. This sample technique was adequate as it allowed the study to acquire important information from participants directly involved in the implementation of multimodal practices. The teachers were engaged in the EFL instruction within the designated schools and had taken part in a professional development initiative that emphasized multimodal literacy. The study was executed in two primary phases. During the initial phase, the teachers engaged in training program focused on multimodal literacy. This program presented fundamental concepts of multimodality and practical strategies for the integration of diverse semiotic modes within classroom instruction. In the course of the training, the teachers devised lesson plans that integrated multimodal resources, which were then executed within their individual classrooms as the second phase. The data were gathered from students subsequent to their engagement with multimodal instruction. This phase concentrated on documenting students' learning preferences regarding the efficacy of various instructional media employed in the EFL classroom context.

Data were collected through a structured questionnaire and open-ended responses. The questionnaire consisted of two sections. The first section contained 26 Likert-scale items examining students' learning preferences and perceptions of multimodal instructional media. The items covered four learning preference categories: visual, auditory, reading/writing, and kinaesthetic. Responses were measured using a five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). The second section consisted of open-ended questions exploring students' experiences and opinions regarding multimodal learning activities and instructional media used in English classrooms.

The questions on the questionnaire were taken from other research projects that looked at multimodal literacy, multimedia learning, and learning interests to make sure they were relevant to the study goals. The instrument was assessed for content validity by two lecturers in English language program of a public university. They assessed the clarity, relevance, and contextual suitability of each item for Indonesian secondary EFL students.

Quantitative data were analyzed using descriptive statistics, including mean, median, mode, standard deviation, frequency, and percentage. The learning preference scores were treated as composite scores because they were calculated by combining several items within each category. Therefore, some mean scores exceeded the original 1–5 Likert scale range. Qualitative data from open-ended responses were analyzed using descriptive thematic analysis.

Descriptive theme analysis was used to look at the qualitative data from the open-ended answers. To get familiar with the data, all of the answers were read over and over again at the start of the analysis. After that, the answers were coded by finding important words, phrases, and statements that connected to see how the students felt about multimodal learning, their favorite teaching materials, how involved they were in class, and their learning experiences. Then, codes that were similar were put into bigger groups and put into recurring themes that showed trends in the answers of all the participants. The themes that were found were then analyzed to support and add to the quantitative results about the alignment between how the students' preferences and they feel about multimodal teaching in the classroom.

RESULTS AND DISCUSSION

Several significant patterns were identified in the findings, including students' learning preferences, media preferences, perceptions of teachers' use of instructional media, and the perceived efficacy of multimodal learning in Indonesian junior high school EFL classrooms. Students expressed a preference for interactive and multimodal learning experiences, especially kinesthetic activities and video-based instructional material. Generally, students also rated teachers' use of instructional media positively, but the suitable between instructional practices and students' preferred learning modes appeared to be realized only partially. Moreover, the results indicated that the blend of several educational media was considered more interesting and useful for learning English than single-mode teaching methods. Detailed results per category are reported in the following sections.

learning styles

The findings indicated that students demonstrated diverse preferences aligned with teacher's use of multimodal teaching in the learning approaches. In this study, data were collected from 60 students using a Likert-scale.

Table 1 *Preference in learning styles*

Learning Styles	Mean	Median	Mode	Sum
Visual	6.39	6.00	6	396
Auditory	6.73	7.00	7	417
Kinesthetic	7.78	8.00	9	482
Reading/Writing	6.65	7.00	6	405

As shown in the table, kinesthetic learning scored the highest mean score ($M=7.78$) indicating that students strongly preferred active learning styles such as role-playing, simulations, and experiments. These preferences indicate that learners gain advantages from active participation in the learning process, as physical activity and experiential methods enhance knowledge internalization and retention. The commonality of kinesthetic preference underscores the necessity of developing interactive learning environments that extend beyond mere passive information reception. This was followed by auditory learning ($M=6.73$), where learners benefited from listening activities, discussions, and oral explanations. Students reported that teacher explanations, participation in discussions, and engagement with audio resources enhanced their comprehension of English. This finding highlights the significance of oral input in the enhancement of listening, comprehension, pronunciation, and speaking abilities. This indicated that numerous students gain advantages from auditory reinforcement in language learning contexts where exposure to spoken English is crucial. The next preference was reading/writing ($M=6.65$). Students who preferred this style typically engaged with reading materials, including books and using writing notes to reinforce their learning. This finding highlights the importance of literacy-based strategies in enhancing comprehension and academic performance, especially for learners who prefer structured, text-driven methods, despite being less dominant than kinesthetic or auditory learning.

Finally, visual learning received the lowest mean score ($M=6.39$). Although comparatively less preferred, it nonetheless served a purpose for certain students who utilized diagrams, charts, and images to clarify complex concepts. The lower score indicates that visual aids only may not be adequate for sustaining engagement; however, they can function effectively as supplementary tools when combined with other instructional methods.

Media preferences

The secondary students' preferences of multimodal learning media preferences is in order to identify which modes of learning resources are most favored by students. The figure below displays the distribution of these preferences.

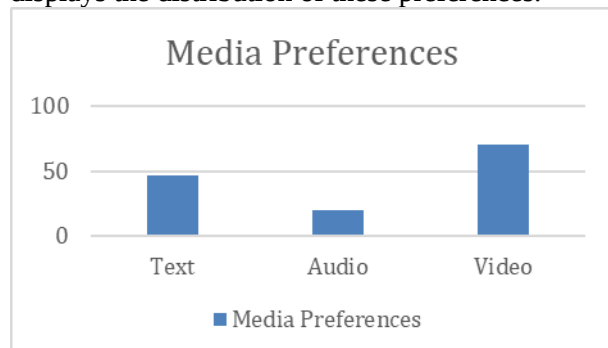


Figure 1. *Media preferences*

Data were collected and analyzed from students' choices on text, audio and video-based materials to find out the sorts of instructional media most liked by secondary EFL students. Figure 1 shows the distribution of students' media preferences. Results indicated that video was the most chosen educational medium with 42 students (70%) selecting it. This shows that students considered video-based learning more interesting and comprehensible, as it integrates visual and audio information concurrently. 28 students (46.7%) chose text-based materials like books and digital

courses. Students found texts valuable for extensive explanations and learning subjects, but often less entertaining than multimedia alternatives. On the other hand, audio-based media such as listening recording and music were the least preferred. Only 12 students (20%) chose audio as main medium of learning. The results indicated that students preferred multimodal teaching resources with visual and aural features rather than single-mode learning medium.

Perceptions of media in english learning

The findings indicate that students generally perceived video-based learning positively. The statement "*Video-based learning helps me understand English material better*" received the highest mean score ($M = 3.63$, $SD = 0.86$), suggesting that students considered video helpful for comprehension. Text-based learning also received relatively positive responses ($M = 3.45$, $SD = 0.90$), while audio-based learning showed lower mean scores ($M = 3.03$ and $M = 3.01$). These results indicate that students tended to perceive multimodal media, particularly videos, as more engaging and supportive for English learning compared to single-mode instructional media. The summary of the results, including mean, standard deviation (SD), and mode for each statement, is presented in the table below.

Table 2 *Students perceptions of media in english learning*

Statement	Mean	SD	Mode
I find it easier to understand English material through texts or reading materials.	3.45	.904	3.00
I feel more comfortable learning English through audio such as recordings or podcasts.	3.03	.943	3.00
Video-based learning helps me to understand English material better.	3.63	.860	3.00
I often choose video-based learning because it is more interesting and easier to follow.	3.27	.963	3.00
I prefer audio because I can listen to it while doing other activities.	3.01	1.076	3.00
Learning English only through text is less engaging compared to video or audio.	3.11	1.14	3.00

The highest score of the statement "*Video-based learning helps me to understand English material better*" reveals that students distinctly appreciate the contribution of video in improving understanding. Videos generally integrate verbal communication with visual elements, gestures, and contextual cues, offering learners various opportunities for comprehension. Furthermore,

videos frequently showcase language in authentic contexts, allowing learners to observe the functioning of vocabulary, grammar, and pronunciation in real-life scenarios (Chen et al., 2024; Sailuddin et al., 2025; Trisanti, N., 2022). This not only strengthens understanding but also enhances engagement by making lessons livelier and more relevant.

Text-based learning received relatively favourable responses, “*I find it easier to understand English material through texts or reading materials*” indicates numerous students perceive text as dependable and accessible medium for structured learning. Texts provide a permanent record of information, that students can review at their own speed, supporting independent study. However, the perception of text alone as “*Less engaging compared to video or audio*” (M=3.11, SD=1.14) underscores the limitations of depending entirely on text-based training. Although texts are essential for developing foundational abilities like reading comprehension and vocabulary development, students recognize that they may lack engagement without additional visual or audio elements. This underscores the need for balance, texts should remain central while being supplemented by different media that enhance interest and engagement.

In contrast, audio-based media was perceived less favorably. The statement “*I feel more comfortable learning English through audio such as recordings or podcasts*” received a lower mean of 3.03 (SD = 0.94), while “*I prefer audio because I can listen to it while doing other activities*” also showed a similar result (M = 3.01, SD = 1.07). Students acknowledge the adaptability of audio, its flexibility and capacity for multitasking, however they typically perceive it as less beneficial for comprehension. This is probably due to audio-only tools lacking the visual aids that assist learners in interpreting meaning. For several student, particularly those still enhancing their listening skills, comprehending spoken English necessitates visual support through gestures, visuals, or written text. The increasing standard deviations for the audio indicate increased diversity in student perceptions, reflecting individual differences in listening proficiency and preference.

These data indicate that students predominantly perceive video as the most effective and interesting medium for English learning, while text remains supplementary resource. Audio, while beneficial in certain contexts, was not highly preferred. This signifies apparent preference towards multimodal resources, especially those simultaneous incorporate visual and audio stimuli, aligning with students' requirements for both comprehension and engagement.

Teachers' use of media

The findings indicate that students generally perceive teachers as attempting to utilize media in alignment with their learning requirements. Students regarded the usage of instructional media by teachers positively. The statement “*The learning media used by teachers helps understanding English material more quickly*” has the highest mean score (M = 3.58, SD = 0.90) and “*Teachers often use learning media that suits my learning style*” (M = 3.53, SD = 0.83). However, the modest mean scores across all categories indicate that the perceived agreement of teachers' multimodal practices with students' chosen learning modalities was only partial and not complete. The results can be seen in the table below.

Table 3 *Teachers' use of media*

Statement	Mean	SD	Mode
The media used by teachers in class suits my learning styles	3.45	.833	3.00
Teachers often use learning media that suits my learning style	3.53	.833	3.00
The learning media used by teachers helps me understand English material more quickly	3.58	.901	3.00

The statement with the highest mean was “*The learning media used by teachers helps me understand English material more quickly*” indicates that student perceptions on teacher-selected media as enhancing learning efficiency. When media are carefully selected and incorporated into classes, they help clarify abstract concepts, reduce processing time, and offer scaffolding for understanding. The positive ratings from students indicate that teachers' use of media is both improving learning outcomes and facilitating the acquisition of new knowledge.

The statement, “*Teachers often use learning media that suits my learning style*” signifies that students recognize teachers' attempts to modify their instruction according to diverse learning preferences, including visual (e.g., visuals, slides, videos), auditory (e.g., audio recordings, discussions), or kinesthetic (e.g., interactive tasks, simulations). This adaptability corresponds with the VARK model of learning styles, which states that learners vary in their preferred methods of receiving and processing information. Teachers seem to be utilizing varied media to address these disparities, thereby promoting diversity in the classroom. Nevertheless, the mean score, while

favorable, remains moderate, suggesting that adaptation may not be consistently or fully beneficial for all learners.

The statement with the lowest rating was “*The media used by teachers in class suits my learning styles*” indicates that while many students feel recognized, there are still discrepancies in the alignment of teachers’ media selections with diverse individual learning preferences. One potential explanation is the practical constraints teachers face in accommodating every student’s style due to limitations in time and resources. Alternatively, some students may lack complete awareness of their own learning preferences, which can affect their perceptions of whether media usage meets their needs. This highlights the intricate nature of differentiated instruction, achieving perfect alignment with each learner’s style remains challenging.

These findings highlight that teacher use of media is generally perceived as supportive and beneficial for students’ learning. However, the moderate mean scores (ranging from 3.45 to 3.58) indicate room for improvement, particularly in ensuring more personalized and consistent alignment between teaching media and students’ varied learning styles.

Effectiveness

To further evaluate the role of multimodal learning in English instruction, students were asked to rate the effectiveness of different media when used by teachers. The findings show that students had a positive opinion of multimodal teaching, especially when different types of media were used together. The statements about using texts, audio, and videos together got the best mean scores ($M = 3.81$), which shows that students thought multimodal instruction helped with both understanding and motivation. When it came to different types of media, videos got the best answer ($M = 3.57$), while audio materials got the worst ($M = 3.04$). Overall, the data show that students thought integrated multimodal resources worked better than single-mode media for learning.

Table 4 *Effectiveness of media in english learning*

Statement	Mean	Mode	SD
I feel helped when teachers use texts in English lessons.	3.53	3.00	0.92
I feel helped when teachers use audio (e.g., listening	3.04	3.00	0.99

exercises or voice recordings).			
I feel helped when teachers use videos in English lessons.	3.57	3.00	0.88
The use of various media (texts, audio, videos) makes it easier for me to understand English material.	3.81	4.00	0.88
I am more motivated to learn when teachers use various media (multimodality) in their lessons.	3.81	3.00	0.90

The table depicts students’ perceptions on the significance of various media in facilitating English learning, specifically focusing on their effectiveness of text, audio, video, and multimodal approaches. The data indicate a similar trend: although each medium possesses value, students expressed the most favorable judgements regarding the integrated usage of numerous media, implying that multimodality is more effective than dependence on a singular channel.

The results indicate that the highest mean scores were associated with statements regarding the integrated usage of various media. Students expressed strong agreement with the statements “*The use of various media (texts, audio, videos) makes it easier for me to understand English material*” with scored ($M = 3.81$, $SD = 0.88$) and “*I am more motivated to learn when teachers use various media (multimodality) in their lessons*” ($M = 3.81$, $SD = 0.90$). This finding highlights the dual advantages of multimodal learning: it improves comprehension and simultaneously increases motivation. This aligns with Mayer (2009) which showed that learning is enhanced when information is delivered through both verbal and visual channels, facilitating more effective meaning-making by students. Furthermore, it aligns with sociocultural perspectives emphasizing that diverse media offer richer contexts and multiple options for engagement, thereby accommodating varied learner needs. In the analysis of individual media effectiveness, video received the highest positive rating. The statement “*I feel helped when teachers use videos in English lessons*” had mean of 3.57 ($SD = 0.88$), this finding indicates that students consider video to be especially beneficial for learning English. The attractiveness of video may be determined by its ability to integrate audio-visual elements, including spoken language, body

language, facial expressions, and contextual environments. These features enhance the practical application of abstract language concepts and offer authentic exposure to English within real-life contexts. Videos serve to connect lessons in the classroom with practical language application, thereby improving comprehension and retention.

Text-based media received positive responses ($M = 3.53$, $SD = 0.92$), indicating its sustained significance in EFL learning. Despite the growing accessibility of digital and interactive tools, students continue to perceive text as reliable and structured source of knowledge. Texts provide stability, clarity, and opportunities for detailed analysis of language forms. This indicates that although multimodality is appreciated, traditional text-based instruction is still crucial, especially for strengthening grammar, vocabulary, and reading comprehension skills.

In contrast, audio-based materials received the lowest mean score ($M = 3.04$, $SD = 0.99$), suggesting that students perceived them as less beneficial compared to the text and video. This may result from multiple factors. Listening comprehension presents challenges, particularly when students encounter unfamiliar accents, speech rates, or pronunciation variations. Secondly, audio alone is deficient in visual cues that facilitate comprehension, which limits learners' ability to process information without additional support. The findings indicate that audio is significant medium for enhancing listening skills, particularly when combined with visual or textual support.

This study's findings highlight the alignment of student's preferences and multimodal teaching practices gives impacts to the transformative potential of multimodal literacy in improving English language teaching and learning in EFL contexts. Students demonstrated significant preferences for multimodal resources, especially video-based materials and kinesthetic learning activities, which enhanced motivation, comprehension, and classroom engagement. These results resonate with broader theoretical and empirical literature, suggesting that multimodality is not only pedagogically effective but also culturally relevant for 21st century learners, especially by engaging students' life worlds and employing culturally responsive pedagogies (Carcamo & Pino, 2025). This research reveals that kinesthetic learning ($M = 7.78$) and video-based resources (70% preference) are the most preferred modalities among students.

The results underscore the active, participative, and context rich approaches that learners perceive as effective. The kinesthetic orientation indicates that students prefer physical involvement and experience learning, along with extensive evidence that interactive activities like role-play and simulations improve knowledge retention (Usuluddin et al., 2024). The significant preference for videos demonstrates learners' tendency towards resources that integrate auditory and visual elements, therefore enhancing the comprehension of input and its relevance to real-world communication.

On the other hand, audio-based content received comparatively lower scores ($M = 3.04$) that requires critical interpretation. Although audio is essential for language acquisition for listening and pronunciation, students may find it less engaging when presented in isolation due to the absence of visual context. The limited preference for audio dual-channel hypothesis, which states that learning is enhance when verbal information is accompanied with visual cues. Other research found that glosses combining video or pictures with definitions are preferred and more effective for vocabulary acquisition that audio-only glosses (Wang & Lee, 2021; Zhang & Zou, 2022). Similarly, although both audio and video enhanced listening skills, video resulted in markedly superior improvements in comprehension, motivation, participation, and retention. Furthermore, research comparing the Video Listening Test (VLT) with the Audio Listening Test (ALT) indicated higher comprehension when visual information is available. This reflects that many 21st century learners, engaged with digital media, anticipate dynamic communication approaches rather than solely static text.

The observed tendencies in student preferences and perceptions correspond with multiple significant theoretical frameworks. Regarding the cognitive Theory of Multimedia Learning it is stated that learners process information more efficiently when it is delivered through both visual-pictorial and auditory-verbal channels (Azar & Tan, 2023; Herlita, 2025; Satar et al., 2023). The significant preference for videos in this study illustrates the dual-channel theory, as students were able to integrate sound, image, and contextual cues to construct meaning more efficiently.

Students recognised their teachers' attempts to integrate media in accordance with learning preferences; however, the moderate scores ($M =$

3.45-3.58) indicate a discrepancy between student expectations and actual classroom practices. This disparity reflects issues observed in Indonesian settings, where despite a curriculum focus on communicative ability, classroom execution frequently remains constrained (De Nadai et al., 2025; He & Suvorov, 2026; Hidayati et al., 2024). Teachers acknowledged the advantages of multimodal texts while also reported obstacles, including infrastructural limitations and insufficient training. Although in-service teachers appreciated multimodal literacy, they necessitated more structured instruction to effectively implement it in their daily practices (Sitokkonni, 2025; Wang et al., 2023). The findings also support the increasing evidence that multimodal literacy is essential for 21st century learning. Multimodal resources support various learning styles and illustrate students' interactions with information in digital environments outside the classroom (Collet et al., 2026; F. Lei et al., 2023; Midgette et al., 2023). This suggests that teachers are encouraged to vary their instructional media by incorporating videos, interactive tasks, and blended modalities to integrate classroom instruction with students' digital practices outside of school. These approaches enhance engagement and facilitate the application of knowledge in practical communication contexts. Finally, these findings indicate that the use of multiple multimodal resources, especially interactive and video-based activities, can enhance instructional alignment and student engagement among EFL teachers in secondary classrooms to address different learning preferences.

CONCLUSION

This study demonstrates that the integration of multimodal literacy, facilitated by teacher training, effectively addresses the challenges present in EFL classrooms, specifically low engagement, limited comprehension, and insufficient contextualized learning. The findings indicate that students exhibit favorable responses to multimodal resources, especially video and kinesthetic activities, whereas teachers gain from structured training that improves their pedagogical awareness and classroom practices. These indicate that multimodal approaches can effectively align curriculum objectives with classroom practices, enhancing communicative competence and learner motivation. The advancement of digital platforms, interactive modules, and ongoing professional development programs is expected to enhance the advantages

of multimodality. Future research may broaden this study's scope and utilize longitudinal designs to assess the sustainable impact of multimodal literacy on English learning practices.

Also, this study presents novel contribution through various dimensions. Initially, it thoroughly investigates the alignment between multimodal teaching practices and student learning preferences, rather than analyzing these constructs in isolation. Secondly, it employs a mixed-method case study approach that synthesizes both teacher and student perspectives, thereby offering a more comprehensive understanding of classroom dynamics. Third, it contextualizes the investigation within the framework of secondary education, an area that continues to be inadequately represented in existing literature. The study ultimately adds to the expanding conversation surrounding multimodal literacy by providing empirical evidence regarding the strategic alignment of multimodal pedagogy with learner diversity, thereby enhancing both engagement and learning outcomes. The study highlights the significance recommendation of professional development in enabling teachers to create and execute multimodal EFL lessons in secondary school level.

REFERENCES

- Amjad, S., Sarwat, S., Asif, A., & Anwar Farooq, M. (2024). ESL learners' perceptions about the effect of language preferences on English learning outcomes. *Journal of Asian Development Studies*, 13(4), 941–956. <https://doi.org/10.62345/jads.2024.13.4.76>
- Azar, A. S., & Tan, N. H. I. (2023). Text presentation or video: Malaysian university students' preferences with synchronous and asynchronous learning. *Education and Information Technologies*, 28(12), 15969–15990. <https://doi.org/10.1007/s10639-023-11796-4>
- Carcamo, B., & Pino, B. (2025). Developing EFL students' multimodal literacy with the use of infographics. *Asian-Pacific Journal of Second and Foreign Language Education*, 10(1), Article 322. <https://doi.org/10.1186/s40862-025-00322-3>
- Chen, B., Zhao, X., & Zhu, Y. (2024). Personalized video summarization by multimodal video understanding. In *International Conference on Information and Knowledge Management, Proceedings*, 1(1). Association for Computing Machinery. <https://doi.org/10.1145/3627673.3680011>
- Collet, V., Araujo, J., Pletcher, B., & Landreth, S. (2026). Affordances of multimodality:

- Introduction to this special issue. *Literacy Research and Instruction*, 65(2), 135–137. <https://doi.org/10.1080/19388071.2026.2628354>
- Creswell, J. W., & Creswell, J. D. (2023). *Research design: Qualitative, quantitative, and mixed methods approaches* (L. Fargotstein, ed.; 6th ed.). SAGE Publications.
- Dayanti, G. R., Ifani, M. Z., Maulana, M., & Ratih, K. (2023). Learning English through hobbies: Exploring students' learning style at English skills classes. *Jurnal Keilmuan Dan Keislaman*, 164–170. <https://doi.org/10.23917/jkk.v2i3.67>
- De Nadai, M., Damianou, A., & Lalmas, M. (2025). Describe what you see with multimodal large language models to enhance video recommendations. In *RecSys2025 - Proceedings of the 19th ACM Conference on Recommender Systems*, 1(1). Association for Computing Machinery. <https://doi.org/10.1145/3705328.3759303>
- de Oliveira, J. M. (2024). Enhancing multimodal literacy through community service learning in higher education. *Frontiers in Education*, 9, Article 1448805. <https://doi.org/10.3389/feduc.2024.1448805>
- Derakhshan, A., Lalli, G. S., & Park, Y. (2026). Unpacking EFL teachers' perceived opportunities and challenges of multimodal AI-driven language testing. *Language Testing in Asia*, 16(1), Article 445. <https://doi.org/10.1186/s40468-026-00445-5>
- Dwi Jayanti, E., & Damayanti, I. L. (2023). Exploring teachers' perceptions of integrating multimodal literacy into English classrooms in Indonesian primary education. *Child Education Journal*, 5(2), 98–109. <https://doi.org/10.33086/cej.v5i2.5240>
- El-Saftawy, E., Latif, A. A. A., ShamsEldeen, A. M., Alghamdi, M. A., Mahfoz, A. M., & Aboulhoda, B. E. (2024). Influence of applying VARK learning styles on enhancing teaching skills: Application of learning theories. *BMC Medical Education*, 24(1), Article 5979. <https://doi.org/10.1186/s12909-024-05979-x>
- Guo, X., Chen, S., & Guo, Y. (2024). Advancing multimodal teaching: A bibliometric and content analysis of trends, influences, and future directions. *Humanities and Social Sciences Communications*, 11(1), 1–22. <https://doi.org/10.1057/s41599-024-04254-0>
- He, S., & Suvorov, R. (2026). Interactive videos in multimodal listening assessments: Examining language learners' perspectives. *TESOL Quarterly*, 0(0), 1–32. <https://doi.org/10.1002/tesq.70099>
- Herlita, A. (2025). The effect of VARK (Visual Auditory Read/Write Kinesthetic) strategy toward students' reading comprehension: A quasi experimental at second grade students' of SMAN 3 Kota Bengkulu. *JKIP: Jurnal Kajian Ilmu Pendidikan*, 6(1), 47–62.
- Hidayati, D., Dharmawan, Y. Y., Prasatyo, B. A., & Luciana, L. (2024). Exploring how translanguaging and multimodal learning improve EFL students' enjoyment and proficiency. *Journal on English as a Foreign Language*, 14(2), 446–475. <https://doi.org/10.23971/jevl.v14i2.8012>
- Humairoh, M. F. N., Widiati, U., Rabbianty, E. N., & Basthomi, Y. (2025). Pre-service English teachers' beliefs about implementing digital multimodal composing in a teaching practicum context. *Journal of Applied Research in Higher Education*, 17(6), 2366–2383. <https://doi.org/10.1108/jarhe-04-2024-0201>
- Indrayani, L., Suprianti, G. A. P., Arthana, I. K. R., & Paramitha, N. L. P. S. (2024). EFL students' learning styles preference and problems encountered in critical reading course. *Jurnal Ilmiah Pendidikan Dan Pembelajaran*, 8(2), 227–234. <https://doi.org/10.23887/jipp.v8i2.68153>
- Lei, F., Cao, Z., Yang, Y., Ding, Y., & Zhang, C. (2023). Learning the user's deeper preferences for multi-modal recommendation systems. *ACM Transactions on Multimedia Computing, Communications and Applications*, 19(3s), Article 3573010. <https://doi.org/10.1145/3573010>
- Lei, Q., & Zhang, C. (2024). Using multimodal resources to design EFL classroom lead-ins—A multimodal pedagogical stylistics perspective. *Linguistics and Education*, 83, Article 101338. <https://doi.org/10.1016/j.linged.2024.101338>
- Li, M., & Dell-Jones, J. (2023). The same topic, different products: Pre-/in-service teachers' linguistic knowledge representation in a multimodal project. *Computers and Composition*, 67Trace, Article 102754. <https://doi.org/10.1016/j.compcom.2023.102754>
- Liu, J. E., & Lim, F. V. (2024). Educational context and teacher beliefs matter: Multimodal literacy in the Greater Bay Area of China. *RELC Journal*, 1–16. <https://doi.org/10.1177/00336882241304136>
- Mark, N. S. J., & Bernadeth, L. F. (2024). Concept analysis of adaptive learning strategy in English language teaching (ALS-ELT). *International Journal of Social Sciences and English Literature*, 8, 45–56. <https://doi.org/10.55220/2576683x.v8.231>
- Midgette, K., Stewart, O., & Gonzales, J. (2023). Teaching with multimodal texts through a critical literacy lens within an English as a foreign language classroom context. *English and American Studies*, 20, 101–109. <https://doi.org/10.15421/382313>

- Norhasanah, N., Yusuf, F. N., & Suherdi, D. (2022). EFL learners' preferences and perspectives on learning styles. *Language Literacy: Journal of Linguistics, Literature, and Language Teaching*, 6(2), 382–399. <https://doi.org/10.30743/ll.v6i2.6172>
- Nurhaliza, & Sari, V. D. A. (2023). Improving eighth graders' vocabulary mastery through students' learning style-VARK-based flashcards. *SAGA: Journal of English Language Teaching and Applied Linguistics*, 4(1), 15–24. <https://doi.org/10.21460/saga.2023.41.164>
- Pikhart, M., Klimova, B., & Al-Obaydi, L. H. (2024). Exploring university students' preferences and satisfaction in utilizing digital tools for foreign language learning. *Frontiers in Education*, 9, 1–9. <https://doi.org/10.3389/feduc.2024.1412377>
- Pillai, M. G., & Khan, F. (2025). Harnessing learning styles and metacognition: A dual framework for academic improvement using the VARK model. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.5091493>
- Pratiwi, D., & Trisanti, N. (2025). EFL learners' preferences for multimodal learning materials at the secondary school level. *Journal of English Language and Education*, 10(3), 630–637. <https://doi.org/10.31004/jele.v10i3.980>
- Pratiwi, W., Sabilah, F., Kurniawati, S., Nurlestari, K., Mujtahidah H. B., Merlinda, R., & Hidayatullah, T. (2024). Merits and demerits of incorporating learning style into EFL classroom instruction: A literature review. *Journal of Languages and Language Teaching*, 12(3), Article 1418. <https://doi.org/10.33394/jollt.v12i3.11222>
- Puspitasari, A., Jailani, M. K., Basthomi, Y., Suharyadi, S., & Zen, E. L. (2025). Multimodal scaffolding in critical reading: Pedagogical insights from Indonesian EFL learners. *Celtic: A Journal of Culture, English Language Teaching, Literature and Linguistics*, 12(2), 924–947. <https://doi.org/10.22219/celtic.v12i2.42721>
- Putri Nur Wandah, Fadillah, J., Firzaly, Z. A., & Asri, S. (2024). Analyzing students' learning styles in listening comprehension. *JET (Journal of English Teaching)*, 10(3), 292–306. <https://doi.org/10.33541/jet.v10i3.6181>
- Quijada, K. A., & Tagle, M. C. (2025). Teaching and learning digital multimodal composing in secondary education: A systematic literature review. *Journal of Computers in Education*. <https://doi.org/10.1007/s40692-025-00371-0>
- Sailuddin, S. P., Rafli, Z., Kamal, M., & Hakim, A. (2025). Audio versus multimodal input: A case study of speech perception among learners in English as a foreign language (EFL) context. *ELS Journal on Interdisciplinary Studies in Humanities*, 8(4), Article 49140. <https://doi.org/10.34050/elsjish.v8i4.49140>
- Sakdiyah, R., Perangin-angin, A., Zein, T. T., & Sinar, T. S. (2026). Multimodal literacy: A visual grammar analysis of Indonesian EFL textbooks in senior high schools. *Journal of Language and Literature*, 5(2), 2615–2628.
- Saputra, B. E., & Dewi, D. N. (2024). VARK learning styles and social media usage in EFL contexts: Insights from tertiary-level students. *English Review: Journal of English Education*, 12(3), 1063–1072. <https://doi.org/10.25134/erjee.v12i3.10783>
- Satar, M., Hauck, M., & Bilki, Z. (2023). Multimodal representation in virtual exchange: A social semiotic approach to critical digital literacy. *Language Learning and Technology*, 27(2), 72–96. <https://doi.org/10.64152/10125/73504>
- Sitokkonni, J. (2025). Bridging perception and practice: How primary teachers adapt instruction based on student learning styles. *ELS Journal on Interdisciplinary Studies in Humanities*, 8(2), Article 45174. <https://doi.org/10.34050/elsjish.v8i2.45174>
- Stewart, O. G. (2023). Using digital media in the classroom as writing platforms for multimodal authoring, publishing, and reflecting. *Computers and Composition*, 67, Article 102764. <https://doi.org/10.1016/j.compcom.2023.102764>
- Sumardi, A. P., & Tahir, M. (2025). Exploring learning style preferences in English speaking class: Students' perceptions from a vocational high school context. *Journal of Vocational Education*, 2, 386–391.
- Tandiana, S. T., Abdullah, F., Andriani, A., Hidayati, A. N., & Rosmala, D. (2023). The mapping multimodal teaching materials for Indonesian EFL students: A need analysis. *Vision: Journal for Language and Foreign Language Learning*, 12(2), 147–168. <https://doi.org/10.21580/vjv12i219990>
- Trisanti, N., et al. (2022). Digital multimodal composing usage in EFL secondary classroom: A study of in-service EFL teachers' perception. *Advances in Social Science, Education and Humanities Research (ASSEHR)*, 375–382.
- Trisanti, N., et al. (2024). University EFL teachers' stance and engagement in digital multimodal composing as multimodal literacy practices. *Journal of Higher Education Language Teaching*, 17(1), 61–70.
- Tsai, K. J. (2025). Concept mapping for enhancing multimodal communication in an EFL classroom. *SAGE Open*, 15(3), 1–16. <https://doi.org/10.1177/21582440251368978>
- Ulu, H., & Cahit, B. (2025). The mediating role of critical thinking disposition in the relationship between multimodal literacy and media analysis skill: Evidence from pre-service teachers. *Education and Information Technologies*, 30,

- 24909–24930. <https://doi.org/10.1007/s10639-025-13697-0>
- Wang, S. P., Chen, Y. L., & Tseng, W. T. (2023). Multimodal instruction and mentor-tutoring in an intensive English program. *SAGE Open*, 13(4), 1–26. <https://doi.org/10.1177/21582440231200151>
- Wang, W., Zheng, L., & Zhang, J. (2026). Effect of online digital storytelling on the comprehension of authentic listening materials and engagement of junior high school EFL learners. *Scientific Reports*, 16(1), 1–15. <https://doi.org/10.1038/s41598-026-36913-4>
- Wiangga, I. A., & Febriyanto, N. I. (2024). The impact of educators' teaching styles on secondary school learners' motivation in English. *Journal of Educational Management and Strategy*, 3(02), 136–144. <https://doi.org/10.57255/jemast.v3i02.361>
- Yi, Y., King, N., & Safriani, A. (2017). Reconceptualizing assessment for digital multimodal literacy. *TESOL Journal*, 8(4), 878–885. <https://doi.org/10.1002/tesj.354>
- Zhu, S., Zhu, X., Li, J., Sun, Y., & Ma, H. (2026). Emotional profiles and their impact on self-efficacy, strategy use, and performance in multimodal reading among primary school students. *Metacognition and Learning*, 21(1), Article 9460. <https://doi.org/10.1007/s11409-026-09460-6>