Available online at www at www.jeltl.org
doi: http://dx.doi.org/10.21462/jeltl.v10i2.1571

Journal of English Language Teaching and Linguistics, 10 (2), 2025

Exploring Senior High School Students' Perceptions of Multimodality in ELT

Endah Luthfiyah[1], Devi Mulatsih[2]
[1, 2]Swadaya Gunung Jati University, Indonesia
endahluthfiyah99@gmail.com, dmulatsih83@gmail.com

Received 19 May 2025| Received in revised form 12 June 2025 | Accepted 27 June 2025

APA Citation:
Luthfiyah, E. & Mulatsih, D. (2025). Exploring Senior High School Students' Perceptions of Multimodality in ELT. *Journal of English Language Teaching and Linguistics, 10(2), 2025,* 261-277.
doi: http://dx.doi.org/10.21462/jeltl.v10.i2.1571

Abstract

The implementation of multimodal approaches in English Language Teaching (ELT) has been increasingly recognized for its ability to enhance student engagement and comprehension. However, few studies have explored how students perceive specific multimodal elements—such as audio/video, visual design, voice and facial expression, gesture, and spatial position—in authentic classroom contexts, particularly at the high school level in Indonesia. This study aims to investigate high school students' perceptions of the use of multimodality in ELT also the benefits and challenges they experience from each element. This study used a qualitative descriptive research design, that data were collected from 30 senior high school students through closed- and open-ended questionnaires. Thematic analysis was used to interpret both quantitative and qualitative responses. The results showed that the most preferred elements of multimodality were audio/video, followed by visual design, voice and facial expression, movement, and spatial position. The benefits felt by learners were that it improved comprehension and helped them to focus in class. However, some challenges were also reported, including technical constraints (e.g., internet instability, unclear audio), fast audio delivery, and excessive teacher gestures, which disrupted students' concentration. This study emphasizes that a carefully designed multimodal approach can create a more inclusive and engaging learning environment. Implications for educators include strategic integration of multimodal elements, professional development in digital literacy and infrastructure improvements. Future research is recommended to explore the direct impact of multimodal strategies on learning outcomes across different contexts.

Keywords: digital learning, ELT, multimodality, students' perception

1. Introduction

The use of multimodality is familiar in language learning. In fact, some recent studies have discussed the implementation of multimodality within the scope of ELT (Aridasarie & Rohmah, 2024; Diyawan, 2024; Khasawneh, 2024). Multimodality relates to the incorporation of several modes used to convey information in a context. The use of multimodality approaches in English language learning can certainly support learners' understanding. Multimodality in education can be integrated with various modes, such as visual, auditory, and kinesthetic elements (Rohi & Nurhayati, 2024). Combining these approaches is an attempt in the classroom to fit with the various characteristics of students concerning their preferred learning environments. It is considered that the use of multimodality in education makes classroom learning more engaging.

Using multimodality in the classroom helps to accommodate the several senses of learners, from sight to hearing to motion. It helps multimodal educational settings to satisfy different personal preferences of students (Bouchey et al., 2021). This approach is also truly relevant in the developmental age of today. The use of media in learning is a tool for teachers (Pulungan, 2021). Many of the learning tools educators employ have evolved from conventional media to media based on digital technologies. The application of digital multimodal in education is rising as digital technology develops (Magnusson & Godhe, 2019). Technology integration in the learning process has a positive impact (Tuzahra et al., 2021). This makes the use of digital media a crucial component of learning activities since it enables teachers and students in conveying and receiving knowledge.

In the setting of language education, the utilization of multimodality-based learning techniques enables more dynamic interaction with the content since they promote language acquisition (Ugli, 2024). Interactive and interesting learning activities can encourage learners' engagement to take an active role in the classroom (Anis & Khan, 2023). Engagement of learners in their education helps them to develop their communication abilities, which is very important for learning language. Including multimodality—that is, modes of sound, image, gesture, and spatial relationships—into ELT helps students to understand the language itself. For pupils in developing communicative competency—the primary focus of language instruction—this multimodal exposure is absolutely vital.

A multimodality approach helps to support current learning ideas emphasizing on customizing education to learners' personality characteristics (Philippe et al., 2020). Grouping students according to their learning styles, such as visual, auditory, and kinesthetic is a follow-up to accommodate learners' interests. To accommodate different learning styles, educators can design learning activities that use multimodality strategies like audio-visual materials, eye-catching presentations, expressive body language, and thoughtful spatial arrangement (Ilmi & Dewi, 2022). Each of these components has the potential to greatly improve students' engagement, language retention, and comprehension, which will improve the efficiency and enjoyment of the learning process. Students with visual learning styles can be accommodated by providing visually stimulating presentations, this aims to help students with visual abilities can break down complex information into parts that are easily digested and remembered by them. Learners with visual and auditory learning styles can be given audio-visual resources to help them recognize and imitate pronunciation from watching and listening. While learning media for students with kinesthetic learning styles can be given lessons that include the teacher's gestures and facial expressions to reinforce meaning and help them remember vocabulary.

Previous studies have extensively acknowledged how multimodal strategies might improve English language acquisition. For example, in a study conducted by Ilmi & Dewi (2022) the results

showed that multimodal elements such as audio/video, gestures, and visual design were positively received by students in Indonesian EFL classrooms. Their study, however, focused on surface-level impressions without investigating to the more fundamental causes of students' different reactions to multimodal use. Similarly, Diyawan's (2024) research emphasized the motivational benefits of using Powtoon in the writing classroom, but the limited sample made her findings less generalizable. Meanwhile, Aridasarie & Rohmah (2024) demonstrated how using traditional puppets with multimodal texts helped speaking development, however their emphasis was on younger pupils in a non-formal education environment. By comparison, Khasawneh (2024) investigated multimodal approaches for Saudi Arabia's dyslexic students, although the demographic and setting were quite different from Indonesian secondary schools.

Although these studies are insightful, they usually show the advantages of multimodality without fully addressing students' different opinions. Despite the virtues of implementing multimodality in learning that can support students' interest and engagement in the teaching and learning process. The diversity of multimodality preferences used by teachers in ELT learning needs to be examined in terms of which multimodality has the most influence in promoting students' interest in learning. Some students may prefer audio or video elements; others may prefer visual design or gesture cues. By understanding which multimodal strategies are most helpful to students from their perspective, educators can better adapt their learning methods to maximize engagement and comprehension. This study aims to explore senior high school students' perceptions of multimodality in ELT. The following research questions have been formulated in this study:

1. How do senior high school students' perceptions of the multimodal elements used in ELT?
2. What are the benefits and challenges perceived by senior secondary students from using multimodal elements in ELT?

2. Literature Review

2.1 Multimodality

Multimodality is a concept that describes the use of more than one mode to convey meaning or messages in an integrated manner. There are various modes that can be used include text, sound, images, gestures, and spatial interactions (Peng, 2019). Using multimodality is a development of the conventional method usually depending just on written or spoken language. Combining several approaches, including sensory and visual aspects, helps communication to be more complete. Learning that implements multimodality can enrich the learning experience by presenting a variety of modalities that allow for deeper understanding, especially since information can be received in many ways (Sharma & Giannakos, 2020). In an educational setting, especially in language instruction, multimodality facilitates a more dynamic and interactive approach that accommodates students with varied learning preferences. The use of multimodality is believed to increase learners' interest in learning because it can motivate them to get involved in learning (Diyawan, 2024). Multimodality focuses on how different modes work together to form meaning that is clearer and easier for students to understand.

2.2 Implementation of Multimodality in ELT

In Indonesia, there is English language learning at the junior secondary education level with the aim that students can learn to communicate using the language. Many approaches are used by teachers to develop skills and support students' understanding of the material in English learning. One of the existing approaches currently used in ELT is the multimodality approach. The implementation of multimodality in ELT has developed along with the improvement of technology

and learning media (Khoiriyah & Ilmi, 2023). Digital tools are a great resource to increase student engagement in language learning (Putri & Damayanti, 2025; Wijewantha, 2021). Educators utilize learning tools to convey knowledge to students in a more engaging, effective, and efficient method. In teaching the English language, multimodality presents an innovative method by integrating numerous elements such as text, imagery, video, and interaction directly. The utilization of various audiovisual materials enables students to interact with language in realistic contexts, so improving their comprehension of language structures and their application in real circumstances. This is in line with Jewitt's (2013) view that emphasizes the importance of various modes of communication in language learning. In addition, activities that actively involve students, such as simulations and role-plays, not only make learning more enjoyable, but also develop comprehensive language skills, as expressed by Kalantzis & Cope (2015) regarding the potential of multimodality in creating more interactive learning.

Implementing multimodality certainly requires careful planning and a deep understanding from the teacher. In addition to designing learning activities that must be adjusted to the characteristics of students, teachers also need to adjust the learning media to the needs and interests of students in choosing learning media (Zainil et al., 2024). The selection of learning modes must be done strategically, considering learning objectives, characteristics of teaching materials, and students' needs and learning styles. The use of various media such as video, audio, images, and text must be harmoniously integrated to create a rich and meaningful learning experience. Teachers also ensure that the multimodality used is effective in delivering material so that it can be well received by students.

In addition, teachers must also have sufficient digital competence to manage various devices used for learning purposes. Teachers with their competence can find ways that can make every student successful in learning (Fadhliyah et al., 2020). This is necessary because the various modes used in learning are already digital-based. Kim (2019) emphasizes the importance of teacher training in this regard, so that they can utilize technology effectively and efficiently. Furthermore, Nagel et al. (2023) argue that the integration of technology in learning is not only limited to the use of tools, but also involves the transformation of the teacher's role into a facilitator in the learning process. Teachers will be able to create meaningful learning experiences for their learners with their digital capabilities. In this context, multimodality can be a powerful tool to create student-centered learning so that they can be actively involved in classroom learning activities.

2.3 Benefits and Challenges in Implementation of Multimodality

Prior studies indicate that the implementation of multimodality in English instruction yields numerous advantages, particularly in enhancing student engagement and motivation. This aligns with the principle of intrinsic motivation, when pupils receive motivation to study from internal sources. This method causes educational resources more engaging and interactive, therefore enhancing student interest and participation in the learning process. By combining various media such as text, images, audio, and video, learning becomes more dynamic and relevant to students' daily lives (Hidayati et al., 2024). Research by Mir (2023), where information presented in visual and verbal forms will be processed by the brain through different pathways, thereby increasing the retention of information that helps in language learning. This is supported by Sitorus et al. (2020) in his study who found that the use of multimodality, such as a combination of text, images, and audio, can improve students' ability to remember vocabulary and understand the context of its use.

However, implementing multimodality in learning also has its own challenges, especially in terms of cognitive load. This is because too much information delivered through multiple modes simultaneously can overwhelm students and actually reduce the effectiveness of learning (Lange &

Costley, 2020). In this context, teachers need to be careful not to cause confusion for students, especially for those who have difficulty processing multiple modes simultaneously. To bridge the gap between implementing multimodal as a medium that helps learning to be effective and a medium that even gives students confusion in the learning process, a study is needed to find what types of multimodal are preferred and effective for student learning in English classes.

3. Research Methods

3.1 Research Design

This study used a descriptive qualitative research design to determine students' perceptions of the use of multimodality in ELT at the senior high school level. This research design was chosen because it focuses on the interpretation and description of findings based on data obtained through questionnaires. This research instrument used a questionnaire adapted from Peng's theory related to elements in multimodal, namely audio/ video, visual design, voice and facial expression, gesture, and spatial position. The results of this study are presented clearly based on the respondents' viewpoints towards integrating multimodality in learning.

3.2 Participants

The participants in this study were 30 students from one of the senior high schools in Indonesia. The participants were selected using purposive sampling technique with the following criteria: (1) high school students who are currently taking English lessons in regular classes; (2) students who have had learning experiences with multimodal approaches in at least one semester; and (3) students who are willing to be respondents voluntarily in this study. The selection of the participants of this study aims to ensure that the participants have relevant experiences with the application of multimodality in ELT so that they can provide responses based on their direct experiences.

3.3 Instruments

Data collection in this study used a questionnaire. There were two types of questionnaires used, namely a closed-ended questionnaire and an open-ended questionnaire. The close-ended questionnaire consisted of 25 statements using a Likert scale to measure the level of student agreement or disagreement with certain statements related to multimodality in learning. The Likert scale used for the close-ended questionnaire has four answer options, namely strongly agree, agree, disagree and strongly disagree. In addition, this study also provided an open-ended questionnaire to explore the participants' deeper understanding of the use of multimodality in ELT. This open-ended questionnaire consisted of four questions. The questionnaire was designed to provide space for students to explain their experiences and perceptions more deeply about the use of multimodality. The use of the open-ended questionnaire aimed to complement the data from the closed-ended questionnaire with exploratory data, so as to provide a more comprehensive portrait of students' perceptions.

3.4 Data Analysis

The data was analyzed descriptively to describe the pattern of students' perceptions of multimodality elements. The data from the closed questionnaire was processed by calculating the percentage of participants' results in each category of answers in each statement. Then, the researcher determined the themes of the statements in the closed questionnaire and presented the percentages based on the resulting themes. Then, the results of the data analysis are presented in tabular form to facilitate interpretation. This was done to provide an overview of the extent to which students agreed or disagreed with various aspects related to the elements of multimodality in ELT.

In addition, data from the open-ended questionnaire were also analyzed by identifying the main themes based on students' answers. Each participant's answer will be categorized according to the relevant theme, namely the benefits they perceive in learning English using multimodal and the challenges they face during learning using multimodality. This data was analyzed to provide an in-depth interpretation of students' perceptions regarding the implementation of multimodality in ELT. This qualitative data is presented in the form of a thematic description and linked to data from the closed questionnaire. The results of this qualitative data will also strengthen the effective multimodal elements used in ELT for learners based on their attractiveness and usefulness.

3.5 Trustworthiness

Data validity in this descriptive qualitative study is maintained through the application of the principle of trustworthiness with four criteria from Lincoln & Guba (1985), namely credibility, dependability, confirmability, and transferability.

First, credibility was established by using a combination of data collection, namely closed and open-ended questions in the questionnaire. Closed questionnaires were used to capture perceptions of multimodal learning and preferences for multimodal elements, while open-ended questionnaires provided space for participants to express their opinions and experiences in an exploratory manner. In addition, the statements and questions on the instrument were adapted from Peng, (2019), so they have a strong conceptual basis and are relevant to multimodality theory.

Second, dependability was maintained through systematic recording of all data collection and analysis procedures. Data collection in this study was conducted online through Google Forms, which facilitated automatic distribution and documentation. In addition, data analysis was conducted using a thematic analysis approach, which identified thematic patterns from students' open-ended questionnaire responses to ensure consistency of results.

Third, confirmability is achieved by presenting data based on the original responses of the participants without the intervention of the researcher's personal opinion. In this article, direct quotes from students' answers to the open-ended questionnaire are presented to support the findings, so that transparency of interpretation can be maintained.

Fourth, transferability is strengthened by presenting a clear contextual description, including the school setting, students' characteristics, and their experiences of multimodal English learning. The participants were upper secondary students who had experienced multimodal learning for at least one semester, so the results of the study can be used as an initial reference for similar educational contexts.

This study guarantees that the data acquired are valid, reliable, and reflect students' opinions really regarding the usage of multimodality in English language learning by employing these four criteria of trustworthiness. Then, this research has obtained ethical approval from the school where it was done. Each participant received an elucidation of the study's objective, assurances of anonymity, and their entitlement to voluntary participation. Furthermore, students were granted the autonomy to withdraw from contribution at any moment without penalties.

4. Results

4.1 Students' Perceptions of the Use of Multimodal in ELT

Based on the results of the questionnaire of 30 students, students' perceptions of multimodality learning strategies in ELT classes have been summarized in the form of a percentage table. The results of 25 questions representing students' views on multimodality integration in ELT

classes are focused on five indicators: audio/ video, visual design, voice and facial expressions, gestures, and spatial placement. The data presented descriptively using the average scores of students' responses in table 1.

Table 1: Analysis of close-ended questionnaire

Elements of Multimodality	Strongly Disagree	Disagree	Agree	Strongly Agree
Audio/ Video	0%	8%	60%	32%
Visual Design	1%	9%	62%	28%
Voice and Facial Expression	1%	10%	53%	36%
Gesture	5%	27%	53%	15%
Spatial Position	5%	35%	48%	12%

This is evidenced by the percentage results from each category of answers which are highest in the "agree" and "strongly agree" categories. The highest percentage is in the visual design element with the "agree" category, where slightly more than sixty percent of total respondents who chose this option. In contrast, the lowest percentage result was found for the audio/video element with no one choosing the "strongly disagree" category. Considering the majority of students gave positive responses and no one strongly disagreed, it was found that there was a strong interest in the audio/video element. This supports the idea that students tend to prefer visual and auditory learning styles.

In addition, the results also showed that almost all respondents favored the use of multimodality in the form of voice and facial expression in ELT, with a total of 89% agreeing and strongly agreeing. Only a small number of students disagreed with the use of this mode. This certainly shows that learners' interest in learning is influenced by the use of the teacher's voice and facial expression mode. However, there are also students who have kinesthetic learning preferences, as seen in the movement and spatial position elements. These elements reached approval levels above fifty percent, although the percentage of disagreement was relatively higher compared to the other elements. This suggests that while the gesture and spatial position elements were appreciated, they were slightly more divisive among the students.

Next, the findings of this study are presented one by one for each multimodality element. First, the results of the closed questionnaire with the theme of multimodal elements, namely audio/video, are presented in Figure 1 below:

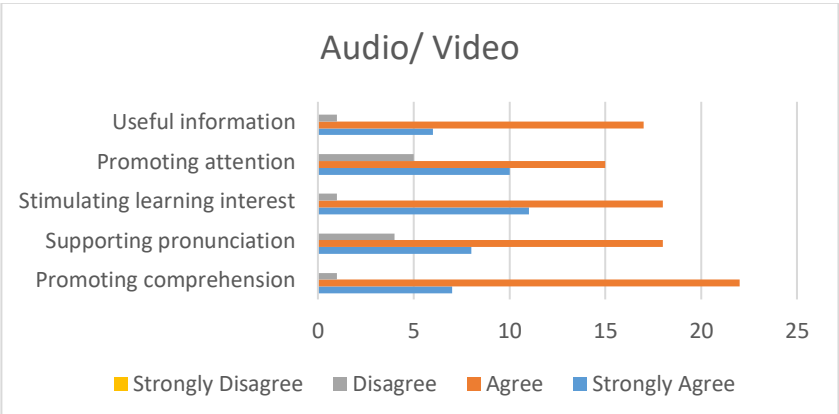

Figure 1: The result of the audio/ video element

Most of the 30 respondents stated that the use of audio/video in ELT is very effective. Almost all students agreed that this element helped to improve their understanding of the material, as reflected by the questionnaire results where 92% of students agreed and strongly agreed. In particular, no students disagreed with this multimodal element. There are 26 of the respondents expressed positive responses to the use of audio/ video that supported their pronunciation. In addition, 96.7% of students gave positive feedback regarding their increased interest. More than three-quarters of students stated that the audio and video aids significantly supported learning by increasing their attention and focus. However, a minority of students, around one fifth of respondent felt that information presented through audio/video was not always fully helpful or relevant to support their understanding and promoting attention. This highlights the need for careful selection and adaptation of audio/video content to better meet students' learning needs.

Second, the outcomes of the data, which pertain to multimodal elements like visual design, are displayed in Figure 2 below.

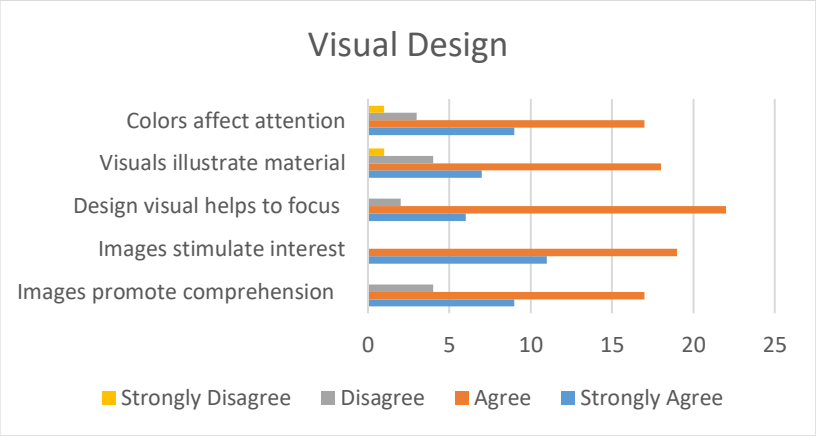

Figure 2: The result of the visual design element

On this element of multimodality, the results showed that almost all respondents gave positive responses to the use of visual design. About 26 out of 30 students stated that pictures help explain the material effectively. The majority of respondents agreed or strongly agreed that the use of pictures in learning can stimulate their interest in the material presented by the teacher. This positive response continues with more than three quarters of students believing that visual design elements, such as color and layout, help them stay focused during learning activities. In addition, most students agree that the use of images, colors, and a well-organized layout make learning more interesting and motivating. Only a small proportion of students, around 3.3% to 10%, felt that visual design had no significant impact on their understanding or focus. These findings suggest that effective visual design can play an important role in improving student engagement and understanding of the material.

Thirdly, Figure 3 below presents the results of a questionnaire that addresses multimodal aspects, including voice and facial expressions.

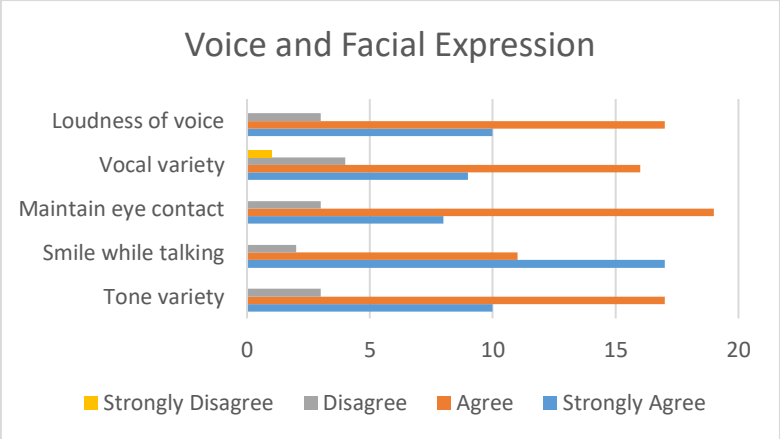

Figure 3: The result of the voice and facial expression element

In this element, most respondents appreciated the teacher's use of variations in tone of voice and facial expressions. Nine-tenths of students agreed or strongly agreed that variations in tone of voice help them to stay focused and engaged during learning activities. In addition, the teacher's smile is very well received by students, as it creates a friendly and comfortable learning atmosphere, as reflected by 93.4% of positive responses. Only a small proportion of students, around 6.7%, disagreed about the importance of teachers smiling when teaching. More than 90% of students felt that the teacher's eye contact helped them feel more cared for and encouraged better interaction. Although some students expressed disagreement regarding the effect of tone of voice and vocal variation on their understanding, the majority of students strongly agreed that tone of voice variation and vocal expression play an important role in increasing their interest and understanding during learning.

Fourth, figure 4 below shows the data collected from the closed questionnaire regarding multimodal elements, specifically gestures:

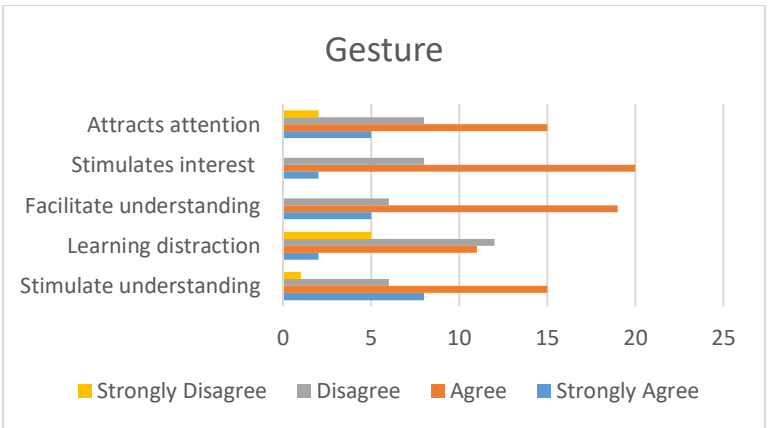

Figure 4: The result of the gesture element

In this section, about 23 out of 30 respondents stated that the teacher's gestures helped improve their understanding of the material. For students, the use of hand gestures to emphasize key points during lesson delivery was helpful in explaining concepts. However, around 43.4% of students stated that excessive gestures can distract their focus during learning activities. Excessive teacher movement was perceived by some students as a potential source of distraction. In addition, more than half of the students responded positively that the use of gestures helped to attract their

attention, making it easier for them to focus on the material presented. These findings suggest that while gestures can significantly support student understanding and engagement, they should be used judiciously to enhance rather than hinder the learning process as they can cause distractions.

Last, figure 5 below illustrates the findings that focusing on multimodal elements, such as spatial position:

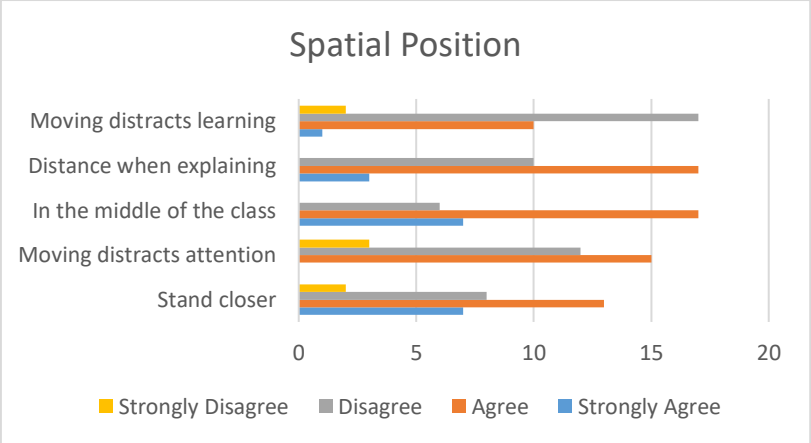

Figure 5: The result of the spatial position element

The findings show that the teacher's standing position in the classroom generated varied responses from the 30 students. Most students, about 66.6%, favored teachers who stand closer to the students or move appropriately during the lesson. In addition, about more than three-quarters of students agreed that when teachers position themselves in the middle of the class it helps to keep learners' attention and makes them feel more involved in the learning process. More than two-thirds of the respondents also emphasized that the proximity of the teacher when explaining the material positively affects their understanding; students prefer if the teacher moves closer rather than maintaining a distant position. However, about half of the students agreed that excessive or undirected movement during a lesson can distract their focus. These findings suggest that teacher movement and positioning are important aspects of classroom dynamics but should be carefully managed to support rather than hinder learning.

In summary, from the analysis of this research instrument, most respondents gave positive responses to the integration of multimodal elements in English language learning, such as audio/ video, visual design, facial expressions, gestures and spatial positioning. Learners feel engaged and interested in learning English with this approach. While most gave positive responses, a small number of students also gave feedback regarding potential distractions or shortcomings of certain elements. This shows that with proper customization, the use of multimodal elements can create a more effective and engaging learning experience for all students.

4.2 Benefits and Challenges of Using a Multimodal Approach

The results of the open-ended questionnaire that completed by 30 senior secondary students, have provided insights into how they perceive the benefits of multimodal in ELT. However, there are also challenges perceived by the respondents in the use of multimodal elements in English language learning. The following is a description of the respondents' views on the use of multimodality in classroom learning. Presented in figure 6 are the advantages of using multimodality in ELT.

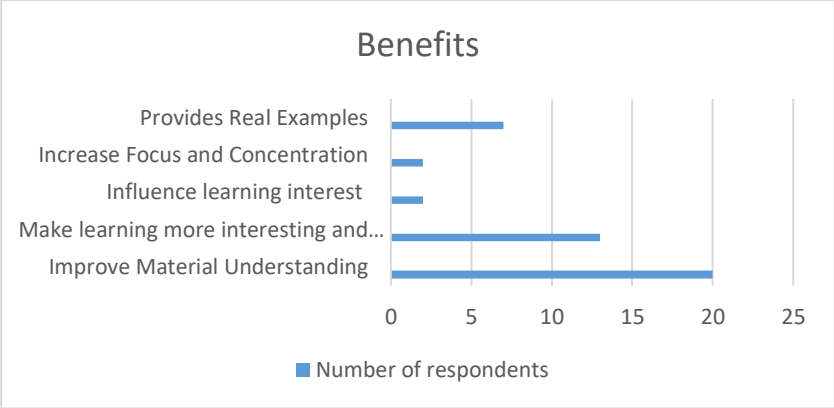

Figure 6: The benefits of using a multimodal approach

Based on the questionnaire results, students revealed their preferred use of multimodality elements as well as the various benefits they felt in ELT. More than half of the respondents revealed that their favorite type of multimodality element is audio/video. One of the respondents, ZA, agreed that "what I like most about using multimodality is video or audio because it is more interesting". Then, the most dominant benefit in using multimodality is to increase understanding of the subject matter (67%). Students stated that the use of media such as video, visual design, audio, gestures and teacher's facial expressions helped them to understand the material more easily. As evidence, RHR said, "With the use of multimodality, students can better understand the learning material because there are examples given from both the teacher's gestures and video/audio". This shows that concrete examples through videos or gestures are highly appreciated because they are able to explain abstract concepts more tangibly.

In addition, almost 50% of students stated that multimodality makes learning more interesting and less boring. Many of them feel that learning becomes more "exciting", "interactive" and "less monotonous" when various multimodal elements are used, as HRS said "by using multimodality, it can make the lesson more interesting and interactive". Videos and visually appealing presentations are considered very informative as well as fun.

Some students, around 23%, also highlighted that multimodality provides real contextualized examples, such as pronunciation demonstrations through audio/video. According to AP "when learning through and using videos, there are English pronunciation sounds with subtitles, this can help to practice English". This certainly helps learners practice their English skills. Another small number of respondents mentioned that multimodality increases their interest in learning and helps them to focus more, especially if accompanied by expressive delivery from the teacher. This is evident in NKH's statement that "multimodality affects my interest in learning English and also I can understand it better through the teacher's expressions". Overall, the findings show that multimodality supports students' diverse learning styles, especially visual and auditory learning styles, thus making English learning more inclusive and effective.

Although most students enjoy using multimodal elements, some challenges are still found. Presented below in Figure 7 are the challenges that learners face in learning English using multimodality.

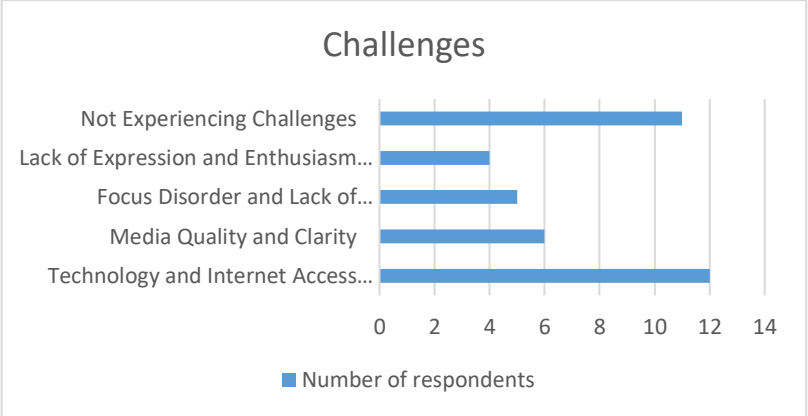

Figure 7: The challenges of using a multimodal approach

The most cited challenge, around 40%, was related to technical constraints, such as unstable internet connections, slow loading videos, or limited access to technological devices. As expressed by RPP "the internet sometimes likes to be slow so that it also disrupts the learning process", a poor internet connection affects the use of multimodality. In addition, according to NKH "videos sometimes take too long when preparing a laptop or projector", this results in a lot of learning time being taken up because of the preparation of displaying videos.

Moreover, the quality of the media used is also an obstacle. Some students complained about unclear audio, uninteresting visuals, or videos that were too fast and did not match the material. This obstacle was expressed by BAW as "the lack of clear video presentation", thus affecting students' understanding of the lesson. Relatedly, one-sixth of students mentioned that they had difficulty understanding the English content, especially when listening to fast or unclear pronunciation in the videos, as KES said "the challenge is that the sound is not clear".

Another challenging aspect is the lack of expression or enthusiasm from teachers. Some students like the multimodal elements of voice and facial expression, so when they feel that if the teacher is not expressive when delivering the material then their enthusiasm for learning also decreases, and the student's mood can be affected. This was expressed by FAR "because my mood depends on the teacher's expression if the teacher's expression is bad then my mood can be bad".

Less than 20% of the respondents mentioned that they had to be more focused when participating in multimodal learning, and others felt limited by their English language skills. MH expressed this challenge as "having to be more focused so that the material can enter the brain" and MAS "because I am not very fluent so when I hear audio/video of people using English, I am a bit confused". However, interestingly, more than 10 students stated that they did not experience significant challenges in using multimodality. They felt comfortable and helped by this approach, and stated that minor obstacles such as signal interference could still be handled easily.

In summary, almost all respondents have a positive perception of the use of multimodality in English language learning. The order of the most preferred multimodality is audio/video, visual design, voice and facial expression, gesture, and spatial position. The most superior use of multimodal elements is in audio/video with the most perceived benefits for students, namely increasing understanding of the material. In addition, teachers' facial expressions and gestures also provide benefits in making learning more interesting. However, there are still challenges faced by respondents, such as technical difficulties and media quality. This affects the learning process and students' understanding of the material. Based on these challenges, it shows that the integration of

multimodal elements needs to be designed more carefully and tailored to students' circumstances. The elimination of these obstacles can help multimodality-based English education to be more inclusive and supportive of students' requirements.

5. Discussion

The results of this study revealed the diverse perceptions of senior high school students towards the application of multimodality in English language learning. In line with the purpose of this study, it was also found how these elements enhance students' engagement and understanding in the classroom learning process. This research shows that students generally have a positive perception of multimodality-based learning strategies, especially on the audio/video element, which is considered as an effective tool in learning. This is evidenced by more than 92% of students agreeing that this element helps them understand the material and be more interested in learning. Students stated that videos help them in understanding the lesson as well as creating a livelier classroom atmosphere so that they are more focused. This finding is reinforced by Mir (2023) who emphasizes the importance of visual and verbal presentation of information to optimize student understanding and retention. Thus, the use of audio/video element helps students in deep understanding of the material and attracts their attention to learning. In addition, the use of audio/video also facilitates learners' interest in the use of technology in learning. This is in line with the development of multimodality in ELT which is now increasingly advanced along with the progress of technology and learning media (Khoiriyah & Ilmi, 2023).

Besides, students also responded positively to the use of visual design elements, including the use of images, color layout, and PowerPoint presentation. Visual design is considered effective in clarifying the material and maintaining student focus. This finding is in line with Magnusson & Godhe (2019) research showing that visual aids can simplify complex concepts. This result is also supported by Dewi et al. (2023) stated that the use of visual media can help students understand the material better. Learners find it helpful to use these multimodal elements in ELT because the visualizations are engaging and encourage learner focus.

Furthermore, students appreciated the intonation of the teacher's voice and facial expressions. Around 90% of students stated that variations in the teacher's tone of voice and facial expressions, such as smiles, created a more comfortable learning atmosphere and improved their concentration. This finding supports Jewitt's (2013) view on the importance of nonverbal communication in a multimodal approach. In fact, some students stated that their mood in learning was greatly influenced by the teacher's expression. This happens because learners perceive a teacher's facial expression when teaching as telling students about how the teacher is feeling at that time (Chalal & Ummto, 2024). Therefore, this suggests that nonverbal communication from teachers plays an important affective role. In addition, students' perceptions of teachers' facial expressions show that multimodality not only impacts cognitively, but also affectively. When teachers are unable to show friendly facial expressions or voice intonation, students feel less motivated. This shows the emergence of negative perceptions from learners towards this multimodal element. This is in line with Yahyazai et al. (2021) which states that facial expressions between teachers and students are a very important part of nonverbal signals in the classroom, this nonverbal communication can build positive interactions in the classroom. Therefore, teachers who are less expressive are considered to hinder student engagement.

However, not all multimodal elements were positively received. There were some multimodal elements that received the opposite response. In the gesture element, although most students felt helped by the use of hand movements to clarify explanations, around 43% of students stated that excessive use of gestures could disrupt their concentration. This finding is reinforced by Afdaliah

(2022), namely that teacher gestures can improve students' understanding of the material, but the use of inappropriate or too many gestures can distract students during the learning process. Therefore, it suggests that the use of gestures should be moderate and purposeful to support the learning process.

Similarly, the spatial position of the teacher during teaching also showed varied responses. Most students consider that teacher movement closer to students can increase attention and interaction. This is in line with Wulandari (2024), where teacher movement in teaching has a positive influence on student engagement and attention in the classroom. However, students' negative perceptions of this multimodal element also emerged, where too frequent and undirected teacher movements were also reported to disrupt students' focus. This result is reinforced by Ilmi & Dewi (2022), that excessive movement from the teacher will disrupt student learning. Therefore, nonverbal aspects such as gestures and spatial placement need to be designed wisely to truly support learning.

In addition to the perceptions of each multimodal element, this study also found some key benefits perceived by students. Students acknowledge that multimodality improves their understanding of the material, makes learning more interesting and less boring, provides concrete contextual examples especially in the aspect of English pronunciation, and increases motivation and learning focus. This finding is strengthened by Diyawan (2024) who state that the use of multimodal can increase learners' interest in learning. Beyond the benefits that learners have experienced from using multimodal before, these multimodal elements also provide advantages in terms of tailoring to the characteristics of each learner. This shows that the multimodal approach is able to accommodate the diversity of students' learning styles. This result supported by Sharma & Giannakos (2020), multimodal learning is able to accommodate the diversity of student learning preferences, both visual, auditory, and kinesthetic, hence making learning more effective and inclusive.

Despite the advantages for learners, difficulties exist in the implementation of multimodality in ELT. Technical obstacles, including unreliable internet connectivity, unclear audio, and challenges in accessing digital devices, constitute the primary issues experienced by students. This issue is in line with Pulungan's (2021) findings that highlight the importance of adequate technological infrastructure in supporting digital-based learning. Facility support is needed so that the use of digital-based modes can run well. This technical barrier occurs due to the lack of availability of facilities such as projectors, internet networks, gadgets/ computers, and speaker. In addition, some students have difficulty understanding the audio content, especially when the speed of speaking in the audio is too fast and unclear or the English vocabulary used is at a high level or not appropriate for the learners' abilities. This is consistent with Lange & Costley (2020) who emphasizes that inappropriate speed of information delivery and selection of material difficulty levels can lead to cognitive overload, hindering comprehension.

These results imply that the effective application of multimodal learning depends not only on the variety of media utilized but also on the teacher's capacity to choose, create, and manage multimodal aspects depending on student characteristics and classroom environment. Teachers need to ensure that the multimodality implemented is in accordance with students' language skills and the availability of adequate technology facilities, as stated by (Magnusson & Godhe, 2019). In implementing multimodal learning, the role of the teacher is also transformed, from being a material deliverer to a facilitator. This is in line with Nagel et al. (2023) who state that technology integration in education extends beyond the utilization of assistive technologies; it also transforms the teacher's position into that of a facilitator in the teaching and learning process. However, to

support this success, the development of teachers' professional competence is also a crucial aspect. This is supported by Tan et al. (2023) who emphasized the importance of teacher training in designing effective multimodal learning. Also, in line with the views of Fadhliyah et al. (2020) and Kim (2019) on the importance of teachers' digital literacy in today's education context. Consequently, educators must possess the capability to choose multimodal aspects that align with students' competencies and interests to provide efficient and accessible education.

Overall, the majority of students have a positive view of multimodality, particularly in relation to audio/video elements, visual design and teacher facial expressions. These multimodal elements have the potential to enhance understanding, stimulate engagement in learning and foster a conducive classroom environment. Nonetheless, negative perceptions remain, such as excessive teaching gestures, uncontrolled movements, and unfriendly intonations or facial expressions. These can disrupt concentration and demotivate students when learning in the classroom. In addition, there are perceived benefits during multimodal English learning such as increased comprehension, motivation, focus, and compatibility with learners' learning styles. However, students also face challenges, especially technical constraints and difficulty understanding audio that is too fast. The success in utilizing multimodality in learning depends on the teacher's ability to select and manage media that are appropriate to the characteristics of the students and the availability of facilities. Therefore, teachers are also required to act as facilitators and have strong digital literacy.

6. Conclusion

The study revealed that students' perceptions of multimodality in ELT varied widely, with most favoring components such as audio/video, visual design, and teacher facial expression and intonation. These elements were perceived to enhance comprehension, engagement and enjoyment of learning. However, the findings also point to barriers, including unfavorable teacher facial expressions, excessive gestures, meaningless movements, and technical and language barriers that can affect students' focus and enthusiasm. The present study indicates that the successful utilization of multimodality in ELT relies not simply on the quantity of modes employed, but also on the connection between multimodal components and the students' necessities, learning preferences, and situations. Effective multimodality is characterized by selective selection, appropriate application, and a focus on enhancing students' cognitive and affective comprehension and participation.

Therefore, educators must possess pedagogical knowledge and technological proficiency in the creation and management of adaptive multimodal resources for learning. Support for schools through training and the deployment of sufficient digital infrastructure is essential. This study has limitations due to its small sample size and singular educational context, thereby restricting generalizability. Further research may encompass a broader and more heterogeneous population and employ experimental techniques to assess the measurable impact of each multimodal component on language learning outcomes.

References

Afdaliah, N. (2022). Teachers' Gestures in EFL Classroom. *Al-Lisan*, *7*(2), 182–197. https://doi.org/10.30603/al.v7i2.2735

Anis, M., & Khan, R. (2023). Integrating Multimodal Approaches in English Language Teaching for Inclusive Education: A Pedagogical Exploration. *Universal Journal of Educational Research*, *2*(3), 2960–3722. www.ujer.orgwww.ujer.org

Aridasarie, Z., & Rohmah, Z. (2024). Enhancing Multimodal Literacy: Puppetry for English Speaking Skills Development. *Cogent Education*, *11*(1).

https://doi.org/10.1080/2331186X.2024.2370754
Bouchey, B., Castek, J., & Thygeson, J. (2021). Multimodal Learning. *Springer*. https://doi.org/10.1007/978-3-030-58948-6_3
Chalal, F. Z., & Ummto, S. Y. (2024). The Role of Teachers' Body Language in the EFL Classroom: A Multimodal Perspective Le Rôle du langage corporel des enseignants dans la classe d'Anglais comme langue étrangère : Une perspective multimodale. *The Algerian Journal of Letters*, *8*(2), 260–273.
Dewi, H. H., Hidayatulloh, S. M. M., Sukarno, Lestari, A. E., Dewi, I. L., & Ciptaningrum, D. S. (2023). English Materials Development for an Undergraduate Communication Study Program: a Need Analysis in Indonesian Context. *LLT Journal: Journal on Language and Language Teaching*, *26*(1), 69–91. https://doi.org/10.24071/llt.v26i1.5208
Diyawan, N. A. A. (2024). Students' Perspectives in The Use of Powtoon as A Multimodal Learning Strategy in Writing Narrative. *InCoLLT*, 253–262.
Fadhliyah, R., Mirizon, S., & Petrus, I. (2020). Professional Development of English Teachers at a State Islamic Senior High School in Palembang. *Indonesian Journal of EFL and Linguistics*, *5*(2), 471. https://doi.org/10.21462/ijefl.v5i2.282
Hidayati, D., Dharmawan, Y. Y., & Prasatyo, B. A. (2024). Exploring How Translanguaging and Multimodal Learning Improve EFL Students ' Enjoyment and Proficiency. *Journal on English as a Foreign Language*, *14*(2), 446–475.
Ilmi, K. A., & Dewi, D. N. (2022). The Integration of Multimodality in EFL Classes: Students' Perception. *Jo-ELT (Journal of English Language Teaching)*, 9(2), 150. https://doi.org/10.33394/jo-elt.v9i2.6343
Jewitt, C. (2013). Multimodal Methods for Researching Digital Technologies. *The SAGE Handbook of Digital Technology Research*, 250–265. https://doi.org/10.4135/9781446282229.n18
Kalantzis, M., & Cope, B. (2015). Learning and New Media. *The SAGE Handbook of Learning*, 373–387. https://doi.org/10.4135/9781473915213.n35
Khasawneh, M. A. S. (2024). Adapting Multisensory Techniques for Dyslexic Learners in English Language Learning: A Case Study Approach. *World Journal of English Language*, 14(5), 553–564. https://doi.org/10.5430/wjel.v14n5p553
Khoiriyah, K., & Ilmi, A. F. (2023). Multimodal Digital Extensive Reading Program: A Reflection On Literacy Movement in Madrasah and Pesantren. *Journal on English as a Foreign Language*, *13*(1), 265–288. https://doi.org/10.23971/jefl.v13i1.6059
Kim, M. S. (2019). Developing A Competency Taxonomy for Teacher Design Knowledge in Technology-Enhanced Learning Environments: A Literature Review. *Research and Practice in Technology Enhanced Learning*, *14*(1). https://doi.org/10.1186/s41039-019-0113-4
Lange, C., & Costley, J. (2020). Improving Online Video Lectures: Learning Challenges Created by Media. *International Journal of Educational Technology in Higher Education*, *17*(1). https://doi.org/10.1186/s41239-020-00190-6
Lincoln, Y.S. and Guba, E. G. (1985). Naturalistic Inquiry. *SAGE*, 289–331. http://dx.doi.org/10.1016/0147-1767(85)90062-8
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
Mir, J. (2023). Impact of Dual Coding Strategy to Enhance Students' Retention of Scientific Concepts in Middle Schools. *Annals of Human and Social Sciences*, 4(IV). https://doi.org/10.35484/ahss.2023(4-iv)63
Nagel, I., Guðmundsdóttir, G. B., & Afdal, H. W. (2023). Teacher Educators' Professional Agency in Facilitating Professional Digital Competence. *Teaching and Teacher Education*, *132*.

https://doi.org/10.1016/j.tate.2023.104238

Peng, J. E. (2019). The Roles Of Multimodal Pedagogic Effects and Classroom Environment in Willingness to Communicate in English. *System*, 82, 161–173. https://doi.org/10.1016/j.system.2019.04.006

Philippe, S., Souchet, A. D., Lameras, P., Petridis, P., Caporal, J., Coldeboeuf, G., & Duzan, H. (2020). Multimodal Teaching, Learning, and Training in Virtual Reality: A Review and Case Study. *Virtual Reality and Intelligent Hardware*, *2*(5), 421–442. https://doi.org/10.1016/j.vrih.2020.07.008

Pulungan, A. H. (2021). The Use of Interactive Learning Media for Teachers in Rural Areas. *Budapest International Research and Critics in Linguistics and Education (BirLE) Journal*, *4*(1), 524–532. https://doi.org/10.33258/birle.v4i1.1705

Putri, Z. F. S. I. L. D. (2025). Pleasure Reading, Better Understanding: A Young EFL Learner's Journey in Digital Extensive Reading. *Journal of Foreign Language Teaching and Learning*, *10*(2). https://journal.umy.ac.id/index.php/FTL/issue/archive/1153

Rohi, M. P., & Nurhayati, L. (2024). Multimodal Learning Strategies in Secondary EFL Education: Insights from Teachers. *Voices of English Language Education Society*, *8*(2), 458–469. https://doi.org/10.29408/veles.v8i2.26546

Sharma, K., & Giannakos, M. (2020). Multimodal Data Capabilities for Learning: What Can Multimodal Data Tell Us About Learning? *British Journal of Educational Technology*, *51*(5), 1450–1484. https://doi.org/10.1111/bjet.12993

Sitorus, L. S., Mardianto, M., & Matsum, H. (2020). Development of Powerpoint-Based Learning Media on Learning Aqeedah Morals. *Budapest International Research and Critics in Linguistics and Education (BirLE) Journal*, *3*(2), 958–964. https://doi.org/10.33258/birle.v3i2.1012

Tan, L., Thomson, R., Koh, J. H. L., & Chik, A. (2023). Teaching Multimodal Literacies with Digital Technologies and Augmented Reality: A Cluster Analysis of Australian Teachers' TPACK. *Sustainability (Switzerland)*, *15*(13). https://doi.org/10.3390/su151310190

Tuzahra, F., Sofendi, S., & Vianty, M. (2021). Technology Integration of the In-Service EFL Teachers: A Study at a Teacher Profession Education Program. *Indonesian Journal of EFL and Linguistics*, *6*(1), 317. https://doi.org/10.21462/ijefl.v6i1.396

Ugli, U. O. K. (2024). The Effectiveness of Technology-Enhanced Language Learning Methods. *Oriental Renaissance: Innovative, Educational, Natural and Social Sciences*, *4*(3), 162-179.

Wijewantha, N. W. S. . (2021). Documentaries as Multimodal Texts to Promote Reading in the Virtual Language Classroom: An Experimental Study. *Journal of English Language Teaching and Linguistics*, *6*(3), 633. https://doi.org/10.21462/jeltl.v6i3.611

Wulandari, S. (2024). Teachers ' Nonverbal Cues in EFL Classrooms : Analyzing the Impact of Movement on Student Engagement and Comprehension. *Journal of English Language Teaching and Literature. 5*(2), 528–535. https://doi.org/10.56185/jelita.v5i2.827

Yahyazai, S. G. S. O. (2021). The Importance of Nonverbal Communication in Teaching. *Ilkogretim Online*, *20*(6), 3043–3051. https://doi.org/10.5958/2249-7137.2020.01930.8

Zainil, M., Helsa, Y., Sutarsih, C., Nisa, S., Sartono, & Suparman. (2024). A Needs Analysis on The Utilization of Learning Management Systems as Blended Learning Media in Elementary School. *Journal of Education and E-Learning Research*, *11*(1), 56–65. https://doi.org/10.20448/jeelr.v11i1.5310