

## STUDENTS' PERCEPTIONS OF YOUTUBE VIDEOS TO ENHANCE ENGLISH VOCABULARY

<sup>1</sup> Maylan Alqaromah, <sup>2</sup> Rulik Setiani

<sup>1,2</sup> Universitas Muhammadiyah Kotabumi, Kotabumi, Indonesia  
<sup>1</sup>alqaromahmalqa@gmail.com\*; <sup>2</sup>rulik.setiani@yahoo.com  
\*corresponding author

### Article History

Received: 2026/04/24

Reviewed: 2026/06/01

Accepted: 2026/06/07

Published: 2026/07/01

### ABSTRACT

The integration of digital media into language learning has reshaped how students acquire English vocabulary beyond the classroom, with YouTube emerging as a widely used platform offering rich audio visual content that supports vocabulary development. Despite its popularity, research on how students perceive and utilize YouTube as a learning medium remains limited, particularly when examined through Walgito's cognitive, affective, and conative dimensions and with a focus on final-year students. This study employed a qualitative descriptive approach to explore students' perceptions of YouTube in vocabulary learning, involving 23 undergraduate students from the English Education Study Program in the preliminary stage, with 14 selected for in-depth interviews through purposive sampling. The adequacy of this sample size was determined by data saturation, as responses consistently revealed recurring patterns without new significant information. Data were collected through semi-structured interviews and analysed thematically to identify patterns in students' experiences. Findings indicate that YouTube facilitates contextual vocabulary comprehension in the cognitive aspect, enhances motivation and enjoyment that strengthen engagement in the affective aspect, and fosters independent learning behaviours such as replaying videos, using subtitles, and taking notes in the conative aspect. Challenges remain, including inconsistent study habits, reliance on personal interests, and potential distractions. In conclusion, YouTube demonstrates strong potential as a medium for vocabulary learning, yet its effectiveness depends on structured and strategic use to maximize benefits.

Keywords: digital learning, EFL, students' perceptions, vocabulary learning, YouTube

### INTRODUCTION

Vocabulary is widely recognized as one of the most fundamental components in learning English as a Foreign Language (EFL). Adequate vocabulary knowledge enables learners to comprehend both spoken and written language and to express ideas effectively in various communicative situations. Without sufficient vocabulary, learners may struggle to understand messages, participate in conversations, or convey their thoughts clearly, even if they possess

adequate grammatical knowledge. As the foundation of language competence, vocabulary directly supports the development of listening, speaking, reading, and writing skills, and its mastery requires a gradual and continuous process. In line with this view, Ibhar (2022) emphasizes that vocabulary development is a crucial step in learning a foreign language, since vocabulary lies at the heart of language acquisition, and without vocabulary mastery, language skills cannot be formed. Nevertheless, although vocabulary development constitutes the core of language acquisition, in practice at the tertiary level this process does not always run smoothly.

Despite its importance, vocabulary learning at the tertiary level still faces several challenges. In many university classrooms, English instruction often relies heavily on textbooks, teacher explanations, and limited classroom activities. While these traditional instructional approaches provide basic knowledge of vocabulary, they may not offer sufficient opportunities for students to encounter vocabulary in authentic communicative contexts. Moreover, limited classroom time often restricts students' exposure to diverse language input and reduces opportunities for meaningful language practice. As a result, students may experience difficulties in expanding their vocabulary knowledge beyond what is explicitly taught in class. Therefore, additional learning resources are needed to provide authentic exposure and support autonomous learning outside the formal classroom environment (Elaoufy, 2025)

The rapid advancement of digital technology has created new opportunities for language learning. In line with this, Hidayat et al., (2024) emphasize that technology provides a significant impact on language acquisition and holds a substantial potential to transform the way languages are learned. One manifestation of this technological development is the integration of digital media platforms. Digital media enables learners to access vast amounts of language input from various sources, allowing them to learn beyond the boundaries of formal classroom instruction. Among the numerous digital platforms available nowadays, YouTube has emerged as one of the most widely used platforms among university students and simultaneously serves as an effective medium for vocabulary learning. Previous studies have emphasized that YouTube, with its multimodal content, significantly supports the development of vocabulary and language skills among EFL students (Abbas et al., 2025). As a global video-sharing platform, YouTube provides a rich variety of English-language content, including instructional videos, podcasts, music videos, movies, vlogs, interviews, documentaries, entertainment videos, and short-form videos. This diversity of content allows students to encounter vocabulary through audiovisual input that combines visual images, spoken language, and contextual situations. Such multimodal exposure can facilitate vocabulary learning by helping students understand meanings through context, observe pronunciation and intonation, and remember vocabulary more effectively (H et al., 2025)

In addition to providing diverse language input, YouTube also offers flexibility and accessibility that support independent learning. Students can access videos anytime and anywhere, replay content when necessary, and select materials that match their learning needs and interests. These features make YouTube a potentially valuable learning tool for language learners who seek additional exposure to English outside the classroom. However, while earlier studies have emphasized these advantages, they often overlook how students' perceptions and usage experiences shape learning outcomes. The preliminary findings of this study suggest that effectiveness is not determined solely by the abundance of content but by how students actively choose, interpret, and utilize digital media based on their interests, motivations, and prior experiences. Consequently, the effectiveness of YouTube as a learning medium is closely related to how students perceive and interpret the platform in relation to their learning activities. Perception plays a significant role because it influences students' attitudes toward a learning medium, their frequency of use, and their willingness to continue using it as a learning resource (Masduqi & Azizah, 2025).

To gain initial insights into students' experiences with YouTube in vocabulary learning, a preliminary questionnaire survey was conducted on February 5, 2025, involving 23 students of the English Education Study Program at Universitas Muhammadiyah Kotabumi. The results of this preliminary survey indicate that most students have used YouTube as a resource for learning English vocabulary. The frequency of use varied among students, ranging from two to three times per month, three to five times per month, and more than five times per month. The types of content accessed were also diverse, including music videos, song lyrics, podcasts, movies, entertainment videos, short-form videos, and formal instructional videos. These preliminary findings suggest that YouTube has become an integral part of students' English learning experiences, although the intensity and patterns of use differ among individuals (Masduqi & Azizah, 2025). Preliminary questionnaire data show that many students consider YouTube videos useful for vocabulary learning. A large proportion of respondents reported that YouTube videos help them expand their vocabulary knowledge. The preliminary data also indicate that listening and speaking skills appear to be more strongly supported in vocabulary learning through YouTube videos. These initial findings provide an important basis for further investigation into students' perceptions, particularly those in their final year, regarding the role of YouTube videos in enhancing English vocabulary (Masduqi & Azizah, 2025)

Interestingly, although some final-year students reported that they currently use YouTube less frequently for vocabulary learning due to academic responsibilities such as thesis writing and graduation preparation, their previous experiences with the platform continue to shape their perceptions. In other words, students' perceptions are not formed solely by their current learning practices but are also influenced by their past experiences with a learning medium. According to Snyder et al., (2015), Recent prior experiences exert a powerful influence on how sensory information is processed and interpreted. This perspective is consistent with the cognitive theory proposed by Bartlett & Kintsch (1995), who explained that human perception and memory operate through schemata, which are organized structures of previous experiences that influence how individuals interpret new information. Gregory (2005) further elaborated that perception operates through a top-down process, meaning that individuals interpret stimuli based on prior knowledge, expectations, and experiences.

According to Bimo Walgito in Madani et al., (2023) explain that perception consists of three aspects. The cognitive aspect is shaped by prior experiences, knowledge, expectations, and ways of thinking, which together form specific beliefs about an object. The affective aspect refers to feelings that may be positive or negative, evaluative in nature, and strongly influenced by cultural values or value systems. This aspect highlights how individuals emotionally respond to situations and how they make judgments based on those emotions. The conative aspect reflects an individual's intention to act in accordance with their attitude. It is closely related to the reasons or goals behind behavior and is manifested in everyday actions that reveal how people perceive particular circumstances or objects. These three aspects serve as the conceptual foundation for the study entitled *Students' Perceptions of YouTube Videos to Enhance English Vocabulary*.

The cognitive aspect, which relates to experience, knowledge, expectations, and ways of thinking, is operationalized in this research through questions that examine how YouTube helps students understand the meaning, usage, and pronunciation of vocabulary, as well as the strategies they employ in learning. The affective aspect, which concerns positive and negative feelings and evaluations influenced by cultural values, is reflected in questions that explore how students feel when learning vocabulary through YouTube and how the platform affects their motivation and confidence. Meanwhile, the conative aspect, which refers to the desire to act in accordance with one's attitude, is measured through questions about the factors influencing students' tendency to use YouTube for vocabulary learning and the practical activities they engage in, such as taking notes, replaying videos, or discussing with peers. Thus, Walgito's

theory of perception not only provides a conceptual framework but also directly shapes the research instrument, enabling a systematic analysis of students' perceptions of YouTube as a medium for vocabulary learning.

Final-year students are considered appropriate participants for this study because they possess relatively extensive and diverse learning experiences. Throughout their academic journey, they have been exposed to various learning strategies, teaching methods, and digital learning media, including YouTube. These accumulated experiences enable them to provide more reflective and critical perspectives regarding the role of YouTube as a learning resource. Their perceptions are therefore not merely descriptive but evaluative, as they are based on their long-term engagement with different learning environments and technologies (Devani et al., 2025)

Although numerous studies have examined the use of YouTube in English language learning, most have concentrated on its effectiveness in improving learning outcomes. Research on students' perceptions has also been conducted, but these studies generally remain limited and rarely employ Walgito's theory of perception, which encompasses cognitive, affective, and conative aspects (Madani et al., 2023). Furthermore, studies that specifically investigate final-year students, who possess broader learning experiences, are still scarce. Thus, the research gap lies in the lack of studies that analyze students' perceptions in depth through Walgito's theoretical framework and within the context of final-year learners. This study contributes uniquely by employing a qualitative approach to explore how final-year students perceive the use of YouTube videos in enhancing English vocabulary (Tahmina, 2023).

Based on the discussion above, this study aims to investigate students' perceptions of the use of YouTube as a medium for enhancing English vocabulary. By examining the cognitive, affective, and conative aspects in depth, the study is expected to provide a more comprehensive understanding of the role of YouTube in supporting vocabulary learning at the tertiary level and to contribute to the development of more effective digital media-based English learning strategies

## METHOD

This study employed a qualitative descriptive approach to explore students' perceptions of using YouTube videos to enhance English vocabulary. A qualitative approach was selected because the purpose of the study was to gain an in-depth understanding of students' experiences, opinions, and interpretations regarding the use of YouTube as a learning medium. Qualitative research focuses on understanding social phenomena from the perspectives of participants and interpreting the meanings they assign to their experiences (Creswell, 2012). Through this approach, the researchers aimed to explore how students perceive the role of YouTube in supporting vocabulary development and how they interpret its contribution to improving their English language skills.

This study was conducted at Universitas Muhammadiyah Kotabumi, specifically in the English Education Study Program. In the pre-study stage, 23 final-year students in the eighth semester of the English Education Department were involved as initial respondents. They completed a screening questionnaire to identify their experiences using YouTube videos as a medium for enhancing English vocabulary learning. From these initial respondents, 14 students who met the inclusion criteria and were willing to participate were selected as the main participants of the study. The participants were selected using purposive sampling. Purposive sampling is a technique commonly used in qualitative research to select participants who possess specific characteristics or experiences relevant to the research objectives (Creswell, 2012). This sampling technique allows the researchers to obtain detailed and meaningful information from individuals who are considered knowledgeable about the phenomenon being studied. The adequacy of this sample size was determined based on the principle of data saturation, as

responses from the 14 participants consistently revealed recurring patterns without generating new significant information. Thus, the number of participants was considered sufficient to capture the range of perceptions relevant to the research objectives.

In this study, the inclusion criteria were established to ensure that the selected participants were aligned with the objectives of the research. The first consideration was the frequency of using YouTube videos per month. Students who accessed YouTube videos at least 3–5 times per month or more than 5 times per month were included, as they were considered to have richer experiences and deeper engagement. This threshold is supported by the theory of processing fluency (Palumbo et al., 2021), which explains that when individuals experience repeated exposure to a stimulus, the process of comprehension becomes easier and faster compared to a novel stimulus. Consistent exposure creates a sense of familiarity, and such familiarity tends to generate positive perceptions.

Therefore, setting the threshold of exposure frequency at 3–5 times per month is considered adequate, as this frequency reflects consistent repeated experiences. In line with the theory of processing fluency, repeated exposure not only produces familiarity and positive perceptions but also provides a foundation for students to evaluate and shape their perceptions regarding how YouTube videos contribute to enhancing English vocabulary. In addition, participants were required to demonstrate their willingness to take part in the interview process. By applying these criteria, the researchers ensured that the participants could provide rich descriptions of their experiences and perceptions regarding the effectiveness of YouTube videos in supporting vocabulary learning.

Data in this study were collected through semi-structured interviews. Semi-structured interviews are a qualitative data collection technique in which the researchers prepares a set of guiding questions while allowing flexibility for participants to elaborate on their responses and for the researchers to ask additional questions during the interview process (Creswell, 2012). This type of interview enables the researchers to explore participants' perspectives in depth while still maintaining a focus on the research objectives. The flexible nature of semi-structured interviews allows participants to express their thoughts, feelings, and experiences more freely compared to structured interviews.

The interview guide was developed based on the research questions and aligned with three key components derived from Walgito's theory, namely Cognitive, Affective, and Conative. which served as the framework to explore students' perceptions of using YouTube to enhance English vocabulary as well as the language skills supported by the platform. This study employed a semi-structured interview approach, meaning that although an interview guide was prepared, the questions were not asked in a rigid or strictly sequential manner, allowing the researchers to develop follow-up questions (probing) that were not predetermined and were adjusted based on participants' responses to gain deeper insights. The interviews were conducted individually to create a comfortable environment for participants to share their experiences openly and to reflect on how they use YouTube in learning English, as well as how it influences their vocabulary acquisition and language skills. With the participants' consent, all interviews were video-recorded to ensure data accuracy and subsequently transcribed verbatim so that every response was documented exactly as expressed, thereby maintaining authenticity and grounding the analysis in the participants' original statements.

The duration of the interviews varied across respondents: DA (5 minutes 28 seconds), CV (5 minutes 31 seconds), PA (3 minutes), APF (6 minutes 13 seconds), M.A.P. (5 minutes 35 seconds), RT (5 minutes 43 seconds), NNF (4 minutes 27 seconds), EDA (3 minutes), MBY (14 minutes 47 seconds), SAA (4 minutes 11 seconds), NA (5 minutes 16 seconds), YK (6 minutes 59 seconds), PHN (5 minutes), and ABK (12 minutes 7 seconds). A total of twelve respondents were interviewed in the classroom at Universitas Muhammadiyah Kotabumi, while two respondents (ABK and MBY) were interviewed at their homes. Prior to data collection,

participants were clearly informed about the purpose of the study, the interview procedures, and the voluntary nature of their involvement, and their consent to be video-recorded was obtained before the interviews commenced. To maintain confidentiality, all personal identifiers were removed, initials were used in the transcripts, and the data were securely stored with access restricted to the researchers. The data were analyzed using the interactive model proposed by Miles et al., (2020) which consists of three main stages: data reduction, data display, and conclusion drawing and verification. These stages occur interactively and continuously throughout the research process.

The first stage was data reduction. Data reduction refers to the process of selecting, simplifying, and organizing the raw data obtained from the interviews. In this stage, the researchers carefully read and reviewed the interview transcripts multiple times to gain a comprehensive understanding of the participants' responses. The researchers then identified and highlighted important statements related to students' perceptions of using YouTube for vocabulary learning. These statements were coded to represent key ideas emerging from the data. Information that was not directly related to the research questions was excluded to maintain focus on the main objectives of the study. Through this process, large amounts of qualitative data were condensed into meaningful units that represented significant aspects of participants' experiences and perceptions.

The second stage was data display. Data display involves organizing and presenting the reduced data in a structured and systematic form to facilitate interpretation. In this stage, the researchers organized the coded data into categories based on similarities in participants' responses. Statements that reflected similar experiences or viewpoints were grouped together to form categories. These categories were then analyzed to identify broader themes that represented common patterns in students' perceptions of using YouTube as a learning tool for vocabulary development. By displaying the data in organized categories and themes, the researchers were able to better understand the relationships among the participants' responses and identify significant findings within the data.

The final stage was conclusion drawing and verification. In this stage, the researchers interpreted the themes that emerged from the data display to draw meaningful conclusions related to the research objectives. The researchers examined how the identified themes reflected students' perceptions of using YouTube to enhance English vocabulary and how the platform contributed to their English language learning. To ensure the credibility and consistency of the findings, the researchers re-examined the interview transcripts and compared them with the identified themes. This verification process helped confirm that the conclusions were strongly supported by the data. The findings were then presented in narrative form and supported by direct quotations from participants to illustrate the themes identified in the study. The purpose of the analysis was not to generalize the results statistically but to interpret and describe participants' experiences in a meaningful way (Creswell, 2012).

## **FINDINGS AND DISCUSSION**

### ***Findings***

Based on the results of interviews with 14 students, an overview was obtained regarding their perceptions of using YouTube as a medium for learning English vocabulary. In general, students' perceptions tend to be positive, although there are variations in experience, strategies, and consistency of learning. YouTube is seen not only as a source of vocabulary, but also as a flexible, personal learning space that suits each student's interests. To provide a clearer picture. The researcher presents the research findings descriptively based on the aspects of perception according to Walgito, which include cognitive, affective, and conative. The result of each aspect can be described below.

***Cognitive Aspect***

The cognitive aspect in this study includes four indicators: experience, knowledge, expectation, and belief.

- **Experience**

In the experience indicator from 14 respondents, the results of the interviews show their learning experiences with vocabulary through YouTube varied, with podcasts, music, subtitles, and documentaries emerging as the most common media. SA emphasized enjoyment in learning, PH highlighted subtitles, RT was accustomed to searching for unfamiliar vocabulary, YK emphasized podcasts, CV highlighted music and subtitles, EDA emphasized music and short videos, APF found YouTube efficient and engaging, DA felt supported by images and sound, NA enjoyed educational channels, MAP emphasized podcasts, NNF preferred lighthearted skits, MBY highlighted music and subtitles, ABK emphasized podcasts and subtitles, and PA emphasized documentaries and foreign channels with note-taking and repetition.

- **Knowledge**

In the knowledge indicator from 14 respondents, the results of the interviews show the vocabulary they acquired was diverse, ranging from idioms and slang to academic terms, phrasal verbs, and everyday expressions. SA and PA mentioned daily vocabulary, PH cited verbs and adverbs, RT found unique vocabulary like eggplant vs aubergine, YK mentioned adjectives such as amazing and unbelievable, CV, NA, and MBY emphasized idioms, EDA highlighted new vocabulary from short videos, APF, NNF, and PA emphasized slang, DA mentioned modern terms like vibe and ghosting, MAP cited idioms like Robbing Peter to pay Paul, MBY added academic terms and pronunciation, and ABK emphasized phrasal verbs.

- **Expectation**

In the expectation indicator from 14 respondents, the results of the interviews show the expectations generally focused on improving fluency, enriching vocabulary, and achieving more natural communication, although specific goals varied. SA, NA, and PA hoped to improve fluency, PH wanted better comprehension, RT expected materials suited to needs, YK wanted advanced vocabulary, CV and EDA aimed to enhance vocabulary and pronunciation, APF wanted to sound like native speakers, DA wanted to use vocabulary in daily communication, MAP wanted to discover new things and apply them in conversation, MBY and NNF wanted to avoid stagnation, and ABK wanted to select content according to personal interests.

- **Belief**

In the belief indicator from 14 respondents, the results of the interviews show that YouTube's effectiveness were consistently positive, with many emphasizing its authenticity, efficiency, and ability to provide real examples of vocabulary use. SA believed enjoyment made learning effective, PH trusted routine practice, RT considered YouTube a relevant source, YK, CV, and ABK believed YouTube enhanced vocabulary, NA and MBY emphasized retention, EDA found new vocabulary through YouTube, APF and MAP highlighted native speakers, DA emphasized viewing intensity, NNF believed vocabulary from YouTube made comments more natural, and PA trusted YouTube to provide real examples.

***Affective Aspect***

The affective aspect consists of two indicators: feelings and emotional evaluation.

- **Feelings**

In the feelings indicator from 14 respondents, the results of the interviews show that feelings were consistently positive, with happiness, enjoyment, comfort, and freedom from pressure dominating their experiences of learning vocabulary through YouTube. SA felt happy and comfortable, PH considered learning a hobby, RT felt free from pressure, YK felt happy and

supported, CV felt happy and suitable, EDA felt happy and interested, APF felt happy and not bored, DA felt happy because of music, NA felt happy and not monotonous, MAP felt calm and comfortable, NNF felt happy and free from pressure, MBY felt happy and supported, ABK felt entertained and not bored, and PA felt happy and relaxed.

- **Emotional Evaluation**

In the emotional indicator from 14 respondents, the results of the interviews show emotional evaluations of YouTube were favorable, highlighting its effectiveness, authenticity, efficiency, and suitability as a learning medium. SA, MAP, NNF, MBY, ABK, and PA considered YouTube effective, PH and NA emphasized enjoyment, YK and MAP highlighted authenticity through native speakers, APF and MBY emphasized efficiency and flexibility, RT gave a score of 8/10, CV considered YouTube appropriate, EDA found it interesting, and DA emphasized it was not boring.

### *Conative Aspect*

The conative aspect includes four indicators: drive or encouragement, reason, goal, and real behavior.

- **Drive or Encouragement**

In the drive or encouragement indicator from 14 respondents, the results of the interviews shows learning drives or encouragement stemmed from personal motivation, curiosity, and interests in podcasts, music, and educational content, as well as encouragement from peers. SA was motivated by visuals, PH by foreign films, RT by content variation, YK and MAP by podcasts, CV and MBY by music, EDA and DA by friends' suggestions, APF and PA by personal motivation, NA by educational content, NNF by curiosity about funny content, and ABK by the need to overcome stagnation.

- **Reason**

In the reason indicator from 14 respondents, the results of the interviews show the reasons for choosing YouTube included its practicality, accessibility, interactivity, and the engaging combination of visuals and audio compared to traditional media. SA and RT emphasized visuals and audio, PH and DA found YouTube more interesting than books, YK highlighted diverse content, CV considered it practical, EDA chose interesting content, NA emphasized interactivity, NNF found it informative and entertaining, ABK emphasized accessibility and variety, MBY considered it completer and more flexible, and PA emphasized flexibility and ease of access.

- **Goal**

In the goal indicator from 14 respondents, the results of the interviews show goals were clear and diverse, ranging from mastering vocabulary and speaking fluently to accessing native speakers and boosting confidence. SA wanted to master vocabulary and speaking, PH aimed to support English studies, RT wanted materials suited to needs, YK wanted advanced vocabulary, CV wanted fluent communication, EDA wanted to improve vocabulary and English skills, APF and MAP emphasized access to native speakers, DA wanted to use vocabulary in daily communication, NA wanted fluency, NNF wanted to avoid stagnation, MBY wanted to recall forgotten vocabulary, ABK wanted to make simple sentences, and PA wanted to increase speaking confidence.

- **Real Behavior**

In the knowledge indicator from 14 respondents, the results of the interviews show real behaviors reflected consistent strategies such as watching with subtitles, listening to music, following educational channels, and using YouTube during free time. SA listened to music, PH watched films with subtitles, RT complemented book learning, YK listened to podcasts, CV routinely watched

with subtitles, EDA used YouTube in free time, APF watched with subtitles, DA listened to music according to mood, NA routinely watched educational channels, MAP noted vocabulary from subtitles, NNF routinely watched English content, MBY used YouTube in free time and listened to music, ABK studied in free time, and PA routinely watched documentaries and foreign channels.

From the overall findings, a consistent pattern emerges that students perceive YouTube as a vocabulary learning medium that is enjoyable, authentic, and effective. The main similarity lies in the positive perception of flexibility and authenticity of the content, while differences appear in learning goals, with some students focusing on daily practical use and others aiming for advanced mastery. Differences are also found in motivation stability, where some students maintain consistency while others are easily distracted, and in learning strategies, where some adopt routine approaches while others rely on situational interest. The key pattern that can be concluded is that YouTube provides an autonomous and rich learning space, but its effectiveness greatly depends on students' self-regulation and the consistency of the strategies they employ. Thus, although the data highlight the strength of YouTube as a supplementary medium, a structured approach is required so that positive perceptions can develop into sustainable and productive vocabulary mastery.

### ***Discussion***

This discussion is based on the theory of perception by Bimo Walgito in Madani et al., (2023), which divides perception into three aspects: cognitive, affective, and conative. In the cognitive aspect, students reported positive experiences in learning vocabulary through YouTube videos, ranging from daily expressions such as "I take a bath" and "I watch television," idioms such as "speak of the devil," adjectives such as "amazing" and "unbelievable," slang such as "turn off" and "back off," to advanced vocabulary. Their expectations to enrich vocabulary and improve speaking and listening skills reflect the cognitive dimension of perception according to Walgito. These findings are consistent with Mayer's (2009) multimedia learning theory, which emphasizes that the integration of visual and verbal input strengthens comprehension. However, when compared with Nation's (2001) vocabulary learning framework, the results reveal a gap between receptive knowledge, that is recognizing words, and productive ability, namely actively using words in communication. Thus, although YouTube provides diverse and authentic input, its effectiveness depends on how learners regulate their practice. This analysis shows both convergences, meaning agreement with the benefits of multimedia learning, and divergence, referring to the persistent gap in productive ability, highlighting the need for structured strategies to complement informal exposure.

In the affective aspect, students show positive emotional responses in the form of feelings of happiness, comfort, and enjoyment. They even describe learning vocabulary through YouTube videos as a hobby. Emotional evaluation consistently considers this medium efficient, enjoyable, authentic, and not boring. This is consistent with Walgito's affective aspect and can be understood through the concept of intrinsic motivation (Ryan & Deci, 2000). Furthermore, Wulandari et al., (2025) confirmed that students expressed positive affective perceptions of YouTube. They felt happy, interested, and not bored when using YouTube as a medium for vocabulary learning. They considered YouTube easy to understand and enjoyable, which increased their learning motivation. This emotional perception strengthens the finding that the affective aspect, comfort and high interest, plays an important role in sustaining vocabulary learning. However, relying too much on interest can also create instability. When interest decreases, learning intensity tends to drop as well. In addition, the nature of YouTube videos as entertainment media can open the possibility of distraction that interferes with learning focus. So, while affective responses are mostly positive, there are also challenges when motivation

depends too heavily on enjoyment. The affective aspect not only supports learning but also shows vulnerabilities that need to be managed to keep focus and consistency.

In the conative aspect, positive perceptions encourage students to perform real behavior. They rewatch videos, use subtitles, note new vocabulary, listen to songs, and follow podcasts. This drive comes from personal motivation, friends' recommendations, and academic needs. The logical reasons given are that YouTube videos are more interesting than books, the content is diverse, and it suits visual learning styles. Their goals are specific, namely to enrich vocabulary, support studies, and improve communication. This is consistent with Walgito's conative aspect and aligns with the concept of self-regulated learning (Zimmerman, 2010). However, not all students show consistent self-regulation. Some are routine, but others depend on mood or situation. Overall, the findings show an ambivalent nature. YouTube videos provide authentic input, motivation, and autonomy, but also present challenges in the form of distraction and lack of structure. Its effectiveness depends on how students manage their learning experiences. The claim about independent learning becomes stronger when it is connected with empirical data, where real behaviors such as rewatching, using subtitles, and noting new vocabulary demonstrate self-regulation strategies, although the level of consistency differs. This is further reinforced by the study of Suli et al., (2022) which shows that Indonesian EFL learners use YouTube to enrich vocabulary in a self-regulated manner, although their consistency varies depending on motivation and situational factors.

## CONCLUSION

This study highlights the importance of understanding students' perceptions of YouTube videos as a medium for learning English vocabulary. The findings confirm positive perceptions across cognitive, affective, and conative aspects, while also revealing challenges in maintaining consistent self-regulation. These results contribute to a broader understanding of how digital platforms can serve as meaningful learning spaces for EFL students. The significance of this study lies in showing that YouTube is not merely a source of entertainment but can function as an authentic and motivating environment for vocabulary acquisition. Pedagogically, the results imply that teachers should integrate YouTube strategically into vocabulary instruction, combining audiovisual input with structured learning activities and guidance in self-regulation. This can help maximize the benefits while minimizing distractions. For future research, it is recommended to investigate comparative studies between YouTube and other digital platforms, or to design classroom interventions that explicitly foster self-regulated learning in digital contexts. Such directions will strengthen the practical application of this study and broaden its relevance to language education.

## REFERENCES

- Abbas, N. F., Awad, Z. E., & Mohaisen, A. K. (2025). YouTube as a Learning Tool Among EFL Learners: A Systematic Review. *Arab World English Journal*, 11, 96–110. <https://doi.org/10.24093/awej/call11.6>
- Bartlett, F. C., & Kintsch, W. (1995). *Remembering: A Study in Experimental and Social Psychology* (2nd ed.). Cambridge University Press. <https://doi.org/10.1017/CBO9780511759185>
- Creswell, J. W. (2012). *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research* (4th ed). Pearson.
- Devani, D., Y, Y. G. S., & Rahmani, E. F. (2025). *Students' Perception Towards the Use of Youtube Videos In Learning English*. 9(2). <https://doi.org/https://doi.org/10.31004/jptam.v9i2.29598>

- Elaoufy, H. (2025). Social Media for Vocabulary Learning: Exploring YouTube's Effectiveness in EFL Contexts – A Study of Video-Based Language Learning. *International Journal of Science and Research Archive*, 15(3), 1122–1136. <https://doi.org/10.30574/ijrsra.2025.15.3.1863>
- Gregory, R. L. (2005). *Eye and Brain: The Psychology of Seeing* (5th ed). Princeton University Press.
- H, K., Sofyan, R. R., & Radhiyani, F. (2025). *EFL Student's Perceptions of Using Youtube Videos for Vocabulary Mastery*. 4(3). <https://doi.org/https://doi.org/10.26858/jtechlp.v4i3,%20Nov.78428>
- Hidayat, S., Rifiyanti, H., & Silvanie, A. (2024). Enhancing English Language Acquisition Through Technology-Based Applications. *Jurnal Pendidikan Bahasa Inggris Proficiency*, 6(1), 23. <https://doi.org/10.32503/proficiency.v6i1.4482>
- Ibhar, M. Z. (2022). *The Impact Of Youtube Media On Efl Students' Vocabulary Acquisition*. 4(6). <https://doi.org/https://doi.org/10.31004/jpdk.v4i6.10625>
- Madani, N. S., Hidayat, M. T., & Hasim, W. (2023). The Students' Perceptions on The Use of Youtube in Learning Vocabulary. *English Education and Applied Linguistics Journal (EEAL Journal)*, 6(3), 122–132. <https://doi.org/10.31980/eeal.v6i3.70>
- Masduqi, H., & Azizah, F. Z. (2025). Students' Perception of The Use of YouTube as A Digital Platform in Vocabulary Mastery. *JLA (Jurnal Lingua Applicata)*, 8(1), 60. <https://doi.org/10.22146/jla.91906>
- Mayer, R. E. (2009). *Multimedia Learning* (2nd ed.). Cambridge University Press. <https://doi.org/10.1017/CBO9780511811678>
- Miles, M. B., Huberman, A. M., & Saldaña, J. (2020). *Qualitative Data Analysis: A Methods Sourcebook* (Fourth edition). SAGE.
- Nation, I. S. P. (2001). *Learning Vocabulary in Another Language* (1st ed.). Cambridge University Press. <https://doi.org/10.1017/CBO9781139524759>
- Palumbo, R., Di Domenico, A., Fairfield, B., & Mammarella, N. (2021). When Twice is Better Than Once: Increased Liking of Repeated Items Influences Memory in Younger and Older Adults. *BMC Psychology*, 9(1), 25. <https://doi.org/10.1186/s40359-021-00531-8>
- Ryan, R. M., & Deci, E. L. (2000). Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Contemporary Educational Psychology*, 25(1), 54–67. <https://doi.org/10.1006/ceps.1999.1020>
- Snyder, J. S., Schwiedrzik, C. M., Vitela, A. D., & Melloni, L. (2015). How Previous Experience Shapes Perception in Different Sensory Modalities. *Frontiers in Human Neuroscience*, 9. <https://doi.org/10.3389/fnhum.2015.00594>
- Suli, M., Kamil, A. B., & Mobit, M. (2022). *YouTube-Assisted Self-Regulated Vocabulary Learning: Indonesian EFL*. 9(1).
- Tahmina, T. (2023). Students' Perception of the Use of Youtube in English Language Learning. *Journal of Languages and Language Teaching*, 11(1), 151. <https://doi.org/10.33394/jollt.v11i1.6883>
- Wulandari, Y., Sain, Y., & Aradhawaty, S. (2025). *Student's Perception About Using Youtube Educational Video in Learning Vocabulary*. 13(1).
- Zimmerman, B. J. (2010). Becoming a Self-Regulated Learner: An Overview. *Theory Into Practice*, 41(2), 64–70. [https://doi.org/10.1207/s15430421tip4102\\_2](https://doi.org/10.1207/s15430421tip4102_2)