

CHAPTER II

REVIEW OF RELATED LITERATURE

A. Review of Theoretical Studies

1. Game-Based Learning

a. Definition of Game-Based Learning

Game-Based Learning (GBL) has emerged as one of the most widely studied and applied instructional approaches in modern education, particularly in English as a Foreign Language (EFL) contexts. Fundamentally, game-based learning refers to the intentional use of games or game-like elements within educational settings to support the achievement of specific learning objectives. This approach differs significantly from entertainment gaming because the game structure in GBL is deliberately designed to achieve academic purposes.

Plass et al. (2020) define game-based learning as an instructional approach in which game design mechanics and structural features are intentionally applied to help students acquire new knowledge, practice and reinforce existing skills, develop higher-order thinking abilities, and prepare for further learning. This definition highlights that GBL is not accidental or incidental, but rather a structured pedagogical strategy grounded in strong theoretical foundations that channels the motivational power of games toward meaningful educational outcomes.

Adipat et al. (2021) further complement this definition by describing game-based learning as an educational approach that employs enjoyable, structured, and purposeful activities to encourage students' active participation. In GBL, games function as cognitive scaffolding that helps students integrate, apply, and test their knowledge through interactive decision-making, thereby positioning students as active participants rather than passive recipients of information.

In the context of English language learning, Reinhardt and Thorne (2020) argue that games naturally support language acquisition by creating conditions for sustained attention, authentic communicative practice, and positive emotional association with the target language. This perspective emphasizes that game-based learning provides an engaging learning environment that encourages active participation and meaningful language use, making it a suitable instructional approach in EFL classrooms.

b. Types of Game-Based Learning: Digital Game-Based Learning and Wordwall

In EFL classrooms, game-based learning is most effectively implemented through digital platforms that allow teachers to create, customize, and deliver interactive game activities connected directly to curricular content. Dimitra et al. (2020) define Digital Game-Based Learning (DGBL) as an innovative learning method that employs computer games and software programs to provide educational value and support teaching, assessment, and evaluation processes.

One of the most widely used DGBL platforms in EFL classrooms today is Wordwall. Wordwall is a web-based educational game platform accessible through the internet that allows teachers to design and implement various interactive learning activities tailored to specific language materials and learning objectives (Maulida et al., 2024). In classroom implementation, teachers access Wordwall through laptops or computers connected to the internet, while students participate using their personal smartphones or available classroom devices, with the game content displayed on a shared projector screen. This setup is practical and feasible in the context of SMK Negeri 1 Panyabungan, where the researcher implemented Wordwall activities during the even semester of the 2025/2026 academic year.

Wordwall provides a wide range of digital activity types, and this study specifically employs three of them based on their theoretical

alignment with vocabulary learning principles. According to Nation (2001), effective vocabulary acquisition requires learners to encounter and use new words through multiple exposures in varied contexts, a condition that all three selected activities are designed to fulfill. The first activity, Match Up, requires students to pair target vocabulary words with their correct meanings or definitions. This activity type directly supports the form-meaning mapping process that is central to vocabulary acquisition, as described by Laufer and Nation (1999), who emphasize that matching tasks build lexical knowledge by requiring students to consciously connect word form with semantic content. The second activity, Quiz, presents multiple-choice questions about the meaning or usage of vocabulary items.

This format activates students' receptive vocabulary knowledge and encourages them to think critically about word use in context, which aligns with Nation's (2001) argument that meaning-focused input tasks are essential for vocabulary consolidation. The third activity, Spin the Wheel, introduces a randomized, spontaneous element in which a vocabulary word is selected at random and students must immediately define it or use it in a sentence. This activity creates a low-anxiety, competitive atmosphere that Maulida et al. (2024) found to be particularly effective in increasing student participation and enjoyment during EFL vocabulary learning sessions. Together, these three activities create a progression from structured recognition tasks to spontaneous production, providing the varied and repeated exposures necessary for meaningful vocabulary retention while simultaneously generating the positive affect and engagement that support interest development.

Empirical evidence also supports the effectiveness of Wordwall in generating positive affective outcomes in EFL classrooms. Sakkir and Jabu (2023) in a pre-experimental study conducted at SMAN 1 Gowa, reported a significant improvement in students' English vocabulary achievement when Wordwall was used as the primary digital game-

based learning platform, with mean scores increasing from 82.34 in the pre-questionnaire to 98.31 in the post-questionnaire. Students also reported that Wordwall improved the classroom atmosphere by making English lessons more interesting and enjoyable, confirming Wordwall's ability to generate positive affective responses among EFL learners. These findings are highly relevant to the present study, which draws on Wordwall's demonstrated capacity to generate positive affective responses and increased student engagement in vocabulary learning contexts. A more detailed review of empirical studies using Wordwall and game-based learning is provided in the Review of Relevant Previous Studies section of this chapter.

c. Principles and Characteristics of Game-Based Learning

For game-based learning to achieve its educational objectives, it must be designed and implemented based on clear principles. Plass et al. (2020) identify four fundamental dimensions that characterize effective GBL: cognitive engagement, affective engagement, motivational-behavioral engagement, and social engagement. These four dimensions ensure that games do not merely provide entertainment, but also genuinely promote deep and meaningful learning.

Plass et al. (2020) also acknowledge that the effectiveness of game-based learning depends heavily on instructional design and classroom context. In vocational school settings such as SMK Negeri 1 Panyabungan, where access to technology may be limited, the implementation of Wordwall must be adapted to the availability of projectors, internet connectivity, and students' devices. In this study, Wordwall activities were designed collaboratively, allowing students to participate through shared or alternating devices to ensure that all learners could engage in the learning process effectively.

Based on the study of Adipat et al. (2021), several core characteristics define effective game-based learning in educational contexts. First, effective GBL contains clear and measurable learning

objectives, ensuring that every game activity is connected to specific content or skill goals. Second, GBL provides immediate and informative feedback, enabling students to understand the consequences of their actions in real time while supporting self-regulation skills. Third, GBL is characterized by progressive challenges, in which the level of task difficulty increases as students advance, thereby maintaining optimal cognitive engagement. Fourth, GBL encourages social interaction and collaboration, motivating students to communicate and negotiate meaning together. Fifth, GBL enhances intrinsic motivation through game mechanics such as points, rewards, and competition, giving students a sense of purpose and achievement throughout the learning process.

d. Advantages and Disadvantages of Game-Based Learning

Game-based learning offers various advantages that have been widely demonstrated in the context of English language learning. Dehghanzadeh et al. (2021), through a systematic review in Computer Assisted Language Learning, found that gamified learning environments consistently produce positive effects on student engagement, motivation, and satisfaction in ESL and EFL contexts. Specifically in English classrooms, GBL increases students' willingness to communicate by reducing emotional barriers, supports vocabulary development through enjoyable practice, builds self-efficacy, and develops critical thinking and collaborative skills (Adipat et al., 2021). In the specific context of Wordwall, Sakkir et al. (2023) found that its interactive and visually appealing design significantly enhanced vocabulary learning outcomes and fostered positive attitudes toward English learning among high school students in Indonesia, confirming that Wordwall produces both cognitive and affective benefits in EFL classrooms.

Despite its many advantages, GBL also presents several challenges. Preparing appropriate game activities can be time-

consuming for teachers. If not managed properly, game sessions may become overly competitive, causing students to shift their focus from learning to winning. In digital GBL contexts, limited device availability or unstable internet connections may also hinder the implementation process. These limitations highlight the importance of careful instructional planning and contextual adaptation when implementing GBL, particularly in rural vocational school settings such as SMK Negeri 1 Panyabungan.

2. Students' Interest in Learning English

a. Definition of Students' Interest

Interest is one of the most influential internal variables in the learning process, widely recognized in educational psychology as a key driver of motivation, attention, and academic achievement. Hidi and Renninger (2006) define interest as a psychological state characterized by focused attention, positive affect, and an increasing engagement with a particular subject, object, or activity over time. This definition identifies interest as a dynamic construct that is simultaneously cognitive and affective, involving both what students think about a subject and how they feel about it.

Renninger and Hidi (2016), in their book *The Power of Interest for Motivation and Engagement*, further elaborate that interest is the single most powerful motivational variable in predicting long-term engagement with a subject. Students who develop genuine interest in a topic are more likely to invest sustained effort, seek out additional information independently, and persist through difficult challenges without external pressure. This is a particularly important insight for English language education, where long-term consistent engagement is necessary for genuine language acquisition.

In the context of learning English as a foreign language, students' interest refers to the degree to which students willingly direct their

attention toward English learning activities, feel positive and enthusiastic about participating in English lessons, and have a genuine curiosity about the English language. When students possess this kind of interest, they transform from passive learners waiting to be taught into active learners who take personal ownership of their language development journey.

b. The Four-Phase Model of Interest Development

The most widely recognized and empirically validated framework for understanding how student interest develops over time is the Four-Phase Model of Interest Development proposed by Hidi and Renninger (2006). This model describes interest not as a fixed trait that students either have or do not have, but as a developmental process that evolves through four sequential, cumulative phases, each characterized by different levels of affect, knowledge, and personal value.

Phase 1: Triggered Situational Interest. This is the initial phase of interest development, in which a student's attention is spontaneously captured by an external stimulus in the learning environment. The trigger can be a novel task format, a surprising piece of information, an engaging activity, or a competitive challenge. At this phase, the student's interest is entirely dependent on the environmental stimulus, and without continuation of that stimulus, interest fades quickly. In the English classroom, a game-based activity is a powerful trigger because it introduces novelty, challenge, and social interaction into what students may normally experience as a routine and predictable subject.

Phase 2: Maintained Situational Interest. If the initial trigger is followed by learning experiences that continue to provide meaningful engagement, personal relevance, and positive affect, the student's interest moves into a more sustained state. At this phase, the student begins to associate the subject with positive experiences and develops a habitual tendency to engage with it. Teachers who consistently apply engaging and varied instructional strategies, including GBL, contribute

significantly to helping students maintain their situational interest across multiple lessons.

Phase 3: Emerging Individual Interest. At this phase, the student's interest begins to take on a personal dimension that extends beyond external stimulation. The student starts to independently seek out information about English, revisit content out of curiosity, and connect the subject to their own identity, aspirations, and personal goals. This phase marks the beginning of intrinsically motivated engagement with English learning.

Phase 4: Well-Developed Individual Interest. This is the most advanced phase, in which interest becomes a stable, self-sustaining, and deeply personal predisposition toward the subject. Students at this phase engage with English learning autonomously and persistently, with a high level of positive emotion, accumulated knowledge, and stored personal value associated with the language (Renninger & Hidi, 2016; Hidi & Renninger, 2006).

This model is the primary theoretical foundation of the present study for two specific reasons. First, the Four-Phase Model explains the mechanism through which Wordwall is expected to produce its effect: the novelty and interactive design of Wordwall activities are theorized to trigger Phase 1 (triggered situational interest), while their consistent use across six instructional meetings is designed to help students reach Phase 2 (maintained situational interest). Second, the theory of interest developed by Hidi and Renninger (2006) and further elaborated by Renninger and Hidi (2016) also provides the basis for the four indicators used in the questionnaire instrument of this study, namely attention, enthusiasm and participation, curiosity, and enjoyment. It is important to clarify that these four indicators are affective and behavioral aspects used to measure students' level of interest at a given point in time; they are not the same as, and should not be equated with, the phases of the Four-Phase Model described above. The phases (triggered situational

interest, maintained situational interest, and emerging individual interest) describe how interest is theorized to develop over time, whereas the four indicators describe what this study observes and measures through the questionnaire.

The relationship between Wordwall-based game activities and students' interest in learning English can be explained through the Four-Phase Model of Interest Development proposed by Hidi and Renninger (2006). Each type of Wordwall activity is theorized to stimulate a different phase of students' interest development. It is important to note that this mapping is a conceptual illustration of a possible mechanism based on the characteristics of each activity type, not a claim that has been empirically tested or confirmed in this study, since this study did not track students' interest longitudinally across the six meetings. The proposed mapping is presented in the following table.

Table 2.1 Relationship between Wordwall Activities and Students' Interest Development Phase (Conceptual Illustration)

Wordwall Activity	Phase of Interest Development	Explanation
Match Up	Triggered Situational Interest	The novelty and interactive nature of matching target vocabulary words with their meanings attract students' initial attention and curiosity toward the learning activity.
Quiz	Maintained Situational Interest	The gradual challenges and immediate feedback provided through quiz activities help sustain students' engagement and interest throughout the learning process.

Spin the Wheel	Emerging Individual Interest	The spontaneity and healthy competition created by this activity encourage students to participate actively and develop a stronger personal involvement in learning English.
----------------	------------------------------	--

The table above illustrates a conceptual possibility of how different Wordwall activities may correspond to different phases of students' interest development, based on the Four-Phase Model proposed by Hidi and Renninger (2006). Through interactive and enjoyable learning experiences, Wordwall activities are expected to trigger and maintain students' situational interest, which may gradually contribute to the development of a more enduring individual interest in learning English. However, because this study measured students' interest only before and after the treatment as a whole, rather than tracking it session by session, this table should be read as a theoretical illustration of a possible mechanism, not as evidence that students actually progressed from one phase to the next.

c. Types of Interest: Situational and Individual Interest

Building on the Four-Phase Model, Hidi and Renninger (2006) distinguish between two broader types of interest that differ in their origin and stability: situational interest and individual interest. This distinction carries a direct implication for the instructional design of this study. If a single game session can only generate triggered situational interest, then the six-meeting Wordwall treatment, delivered within this research's overall eight-session schedule, is theoretically calculated to give students enough repeated positive experiences to move toward a more sustained pattern of interest in English learning.

Situational interest is a temporary psychological state triggered by specific features of the learning environment. It is characterized by focused attention, curiosity, and positive emotion directed at a specific

task or activity in the moment. Situational interest is external in origin, meaning it depends on what the teacher provides. In the English classroom, a game activity, an interesting video, or a collaborative task can all trigger situational interest. If teachers fail to provide such stimuli regularly, situational interest will not develop. Research has shown that situational interest significantly predicts students' learning behavior, including effort, persistence, and participation (Renninger and Hidi, 2016).

Individual interest is a stable, long-term predisposition toward a specific content area or activity that resides within the learner rather than being dependent on external conditions. Students with individual interest in English seek out language learning opportunities independently, maintain positive emotional connections to the subject, and persist in their engagement even in the absence of external reinforcement. Individual interest is closely associated with intrinsic motivation, autonomy, and deep learning, all of which are essential for genuine language proficiency development (Hidi and Renninger, 2006).

The pedagogical implication of this distinction is that teachers have the responsibility and the power to develop students' individual interest by consistently providing high-quality, engaging, and personally meaningful learning experiences that trigger and maintain situational interest over time.

d. Indicators of Students' Interest in Learning English

To make the concept of students' interest in learning English measurable and observable for research purposes, this study identifies four key indicators of interest — attention, enthusiasm and participation, curiosity, and enjoyment — adapted from the dimensions of interest described by Renninger and Hidi (2016). These four indicators are aspects used to measure the level of students' interest at a single point in time; they are distinct from, and should not be treated as equivalent to, the developmental phases of the Four-Phase Model (triggered

situational interest, maintained situational interest, and emerging individual interest) discussed in the previous section. These indicators were measured using a 20-item Likert-scale questionnaire administered before and after the treatment.

Attention refers to the degree to which students voluntarily and consistently direct their cognitive focus toward English learning activities during classroom sessions. According to Renninger and Hidi (2016), focused attention is the most basic and foundational indicator of interest because all phases of interest development begin with a learner directing their cognitive resources toward a specific subject or activity. Students with a high level of attention in English lessons actively listen, follow the teacher's instructions, and concentrate on tasks without being easily distracted by irrelevant stimuli.

Enthusiasm and Participation refers to the level of eagerness and proactiveness students demonstrate in engaging with English learning activities. Hidi and Renninger (2006) describe this dimension as the affective and behavioral component of interest, in which students who have developed situational or individual interest show visibly positive energy and take initiative in learning tasks. Students with high enthusiasm actively volunteer answers, participate in discussions spontaneously, and display genuine engagement during English lessons.

Curiosity refers to students' intrinsic drive to explore English language content beyond the boundaries of the prescribed classroom curriculum. Renninger and Hidi (2016) identify this dimension as the knowledge-seeking component of interest, in which learners who are genuinely interested in a subject begin to independently seek out additional information beyond formal classroom instruction. Observable behaviors include voluntary question-asking during class, independent exploration of unfamiliar words, and self-initiated practice of English outside school hours.

Enjoyment refers to the positive emotional experience students associate with learning English. According to Renninger and Hidi (2016), positive affect is a central and non-negotiable element of interest at every phase of its development, since interest cannot be sustained without an emotional component that makes engagement feel rewarding rather than burdensome. Students who genuinely enjoy English lessons experience pleasure in the learning process, view classroom sessions positively, and maintain a generally favorable attitude toward the subject.

The four indicators were selected because they are relevant to the characteristics of Wordwall-based learning activities and can be measured through the questionnaire administered before and after the treatment. These indicators are also consistent with previous studies on gamified learning, which highlight attention, participation, curiosity, and enjoyment as key dimensions of student engagement in digital game-based instruction (Dehghanzadeh et al., 2021; Dindar et al., 2021). The questionnaire consists of 20 items, including 10 positive and 10 negative statements, which were designed to ensure response balance, reduce response bias, and avoid respondent fatigue during data collection.

e. Factors Influencing Students' Interest in Learning English

Students' interest in learning English is not formed in isolation but is shaped by a combination of internal and external factors. Understanding these factors is important because it allows researchers and teachers to design interventions that target the actual roots of low interest rather than addressing only its surface symptoms.

Among the internal factors, Hidi and Renninger (2006) emphasize that prior knowledge and prior positive experiences with a subject are the two most influential internal predictors of interest development. Students who have previously experienced success or enjoyment in English tend to associate the subject with positive

emotions, making them more likely to engage in the future. In contrast, students who have repeatedly encountered difficulty or embarrassment in English lessons often develop avoidance behaviors and reduced self-efficacy, both of which suppress the formation of situational interest.

Among the external factors, the quality of the instructional approach is one of the most influential variables. Muslim et al. (2020), in a study involving Indonesian senior high school students, found that low investment in English learning was strongly associated with a school environment that failed to provide engaging, motivating, and personally relevant learning experiences. The study specifically identified teacher-centered and monotonous instructional methods as key factors that suppress students' interest in learning English. In addition, the quality and variety of learning materials, the classroom atmosphere, the nature of peer interaction, and the teacher's ability to create an enjoyable learning environment all directly influence whether students develop or sustain their interest in English.

3. The Relationship between Game-Based Learning and Students' Interest

The connection between game-based learning and students' interest in English learning is both theoretically coherent and empirically supported. Theoretically, GBL is designed to produce precisely the conditions that the Four-Phase Model of Interest Development (Hidi and Renninger, 2006) identifies as necessary for triggering and sustaining student interest. The novelty and challenge of game activities trigger situational interest (Phase 1). The consistent application of GBL across multiple sessions maintains that situational interest (Phase 2). Over time, as students come to associate English with positive and enjoyable experiences, the conditions for emerging individual interest (Phase 3) are also created.

Plass et al. (2020) explicitly argue that games engage learners on four interconnected levels: cognitively, affectively, motivationally-

behaviorally, and socially. Each of these levels maps directly onto the dimensions of interest identified by Hidi and Renninger (2006). Cognitive engagement corresponds to the focused attention dimension of interest. Affective engagement corresponds to the positive emotion dimension. Motivational-behavioral engagement corresponds to the value and effort dimension. Social engagement corresponds to the collaborative aspect of interest, which involves sharing enthusiasm and knowledge with peers.

Empirical support for the theoretical connection between GBL and student interest is substantial. The specific studies that provide this empirical foundation are reviewed in detail in Section B (Review of Relevant Previous Studies) of this chapter, which examines three prior studies directly relevant to the present research. Dehghanzadeh et al. (2021) found in their systematic review that gamified environments consistently produced positive outcomes in engagement, motivation, and satisfaction among language learners. Deep et al. (2024) further demonstrated that gamification creates a stimulating, low-anxiety environment that enhances students' enjoyment of English learning and reduces the psychological barriers that suppress active participation.

Based on this theoretical and empirical convergence, this study argues that game-based learning, when implemented deliberately and consistently in the English classroom, has the capacity to produce a significant positive effect on students' interest in learning English.

4. Wordwall and Students' Interest: Theoretical Connection

Theoretically, Plass et al. (2020) argue that games engage learners simultaneously at cognitive, affective, motivational-behavioral, and social levels. Each of these levels corresponds directly to the dimensions of interest identified by Renninger and Hidi (2016), namely focused attention, positive emotion, personal engagement, and social value. When students participate in Match Up, Quiz, and Spin the Wheel activities through Wordwall, the novelty and challenge of the game format function as external stimuli that trigger situational interest as described by Hidi and

Renninger (2006). If these positive learning experiences are consistently maintained across multiple instructional sessions, situational interest can gradually develop into a more sustained and personally meaningful engagement with English learning.

B. Review of Relevant Previous Studies

To situate the current study within the existing body of knowledge and to identify both the empirical support and the research gap that this study addresses, this section reviews three relevant prior studies that share significant similarities with the present research in terms of platform, method, context, or dependent variable.

1. **Sakkir, G., Azis, N., & Jabu, B. (2023). Using the digital game Wordwall to enhance EFL students' vocabulary mastery. *Journal of Educational Science and Technology (EST)*, 9(3), 246. <https://doi.org/10.26858/est.v9i3.56966>**

This study investigated the effectiveness of Wordwall as a digital game-based learning platform in improving English vocabulary mastery among high school students at SMAN 1 Gowa, South Sulawesi, Indonesia. The study employed a pre-experimental approach using a one-group pretest-posttest design involving 35 students from Class XI MIPA 3. Students were taught English vocabulary through Wordwall game activities conducted over several learning sessions. Statistical analysis using a paired-sample t-test revealed a significant improvement, with the mean score increasing from 82.34 in the pre-questionnaire to 98.31 in the post-questionnaire, with a significance value below 0.05. In addition to improving vocabulary mastery, students also reported that Wordwall enhanced the classroom atmosphere, making English lessons more interesting and enjoyable.

This study is the most relevant to the present research because it shares similarities in the platform used (Wordwall), the research design (pre-experimental with a one-group pretest-posttest design), the national context (Indonesia), and the educational level (high school students). The main

difference, which also becomes the research gap addressed in the present study, is that Sakkir et al. (2023) focused on cognitive outcomes in the form of vocabulary mastery scores, whereas the present study specifically examines affective outcomes, namely students' interest in learning English through a Wordwall-based approach applied to English vocabulary learning at the vocational high school level in Panyabungan.

2. **Dindar, M., Ren, L., & Järvenoja, H. (2021). An experimental study on the effects of gamified cooperation and competition on English vocabulary learning. *British Journal of Educational Technology*, 52(1), 142–159. <https://doi.org/10.1111/bjet.12977>**

This study employed a controlled experimental design to examine how gamified cooperative and competitive conditions affected English vocabulary learning outcomes and student engagement. The findings showed that students in gamified learning conditions achieved significantly greater improvements in vocabulary mastery compared to students in conventional instructional conditions. In addition, gamified conditions were associated with higher levels of student engagement and willingness to participate in language tasks.

This study provides strong empirical support for the argument that game-based elements in English learning can produce meaningful improvements in both cognitive and affective outcomes. However, this study was conducted outside Indonesia, was not carried out in a vocational school context, did not use Wordwall, and did not specifically measure students' interest as a construct distinct from student engagement. These gaps become the focus addressed in the present study.

3. **Deep, D., Ghosh, A., Gaither, M., & Koptelov, I. (2024). Gamification techniques and the impact on motivation, engagement, and learning outcomes in ESL students. *RAIS Journal for Social Sciences*, 8(2). <https://journal.rais.education/index.php/raiss/article/view/233>**

This study investigated the impact of gamification techniques on motivation, engagement, and learning outcomes among ESL students in a

higher education context. The findings revealed that gamification was able to create a stimulating, relaxed, and enjoyable learning environment that reduced language anxiety, increased students' enthusiasm for participating in English activities, and enhanced overall learning motivation. The researchers emphasized that game-based methods allow students to practice English in a low-pressure environment, which is essential for developing positive affect and sustained interest in the language.

This study supports the main argument of the present research that game-based learning can generate positive affective responses among English learners. The research gap lies in the research context, as Deep et al. (2024) focused on ESL students in higher education, whereas the present study focuses on EFL students in an Indonesian vocational high school using Wordwall as a specific and contextualized learning platform.

A cross-analysis of the three studies above reveals a consistent pattern that game-based learning, whether implemented through general gamification (Deep et al., 2024), gamified competitive learning conditions (Dindar et al., 2021), or specifically through Wordwall (Sakkir et al., 2023), consistently produces positive effects on students' engagement, motivation, and learning outcomes in English language learning contexts. This body of evidence strengthens the theoretical predictions proposed by Plass et al. (2020) as well as the interest development framework developed by Hidi and Renninger (2006).

However, a specific and unaddressed research gap remains evident. None of the three studies investigated the effect of Wordwall on students' interest in learning English within the context of vocational high schools (SMK) in Indonesia. Sakkir et al. (2023) used Wordwall in Indonesia but focused on vocabulary mastery as the learning outcome. Dindar et al. (2021) examined student engagement in English learning, but the study was not conducted in Indonesia and did not use Wordwall. Meanwhile, Deep et al. (2024) investigated affective responses toward gamification, but within an ESL higher education context. No previous study has simultaneously

combined: (1) the Wordwall platform, (2) students' interest as the dependent variable, (3) a vocational high school context, and (4) the Indonesian EFL setting. Therefore, the present study is specifically designed to fill this research gap.

C. Conceptual Framework

The conceptual framework of this study explains the theoretical and logical relationship between the two main variables, namely Wordwall game-based learning as the independent variable (Variable X) and students' interest in learning English as the dependent variable (Variable Y).

Variable X is based on the theoretical frameworks proposed by Plass et al. (2020) and Adipat et al. (2021). Wordwall is understood as a structured instructional treatment in which three types of activities, namely Match Up, Quiz, and Spin the Wheel, are used to engage 27 students of Grade XI AKL in learning English vocabulary in a cognitively active, affectively positive, and socially interactive manner across six instructional meetings during the odd semester of the 2026/2027 academic year.

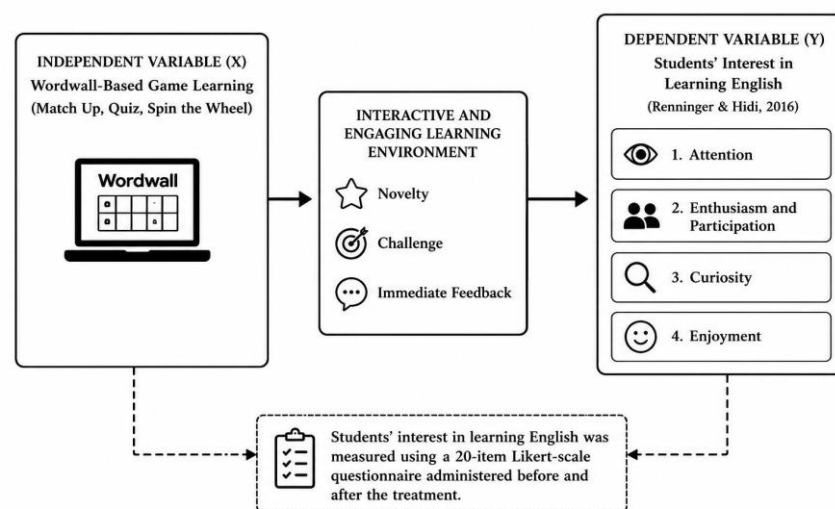
Variable Y is based on the theoretical frameworks developed by Hidi and Renninger (2006) as well as Renninger and Hidi (2016). Students' interest in learning English is understood as a dynamic psychological construct measured through four observable indicators: (1) Attention, (2) Enthusiasm and Participation, (3) Curiosity, and (4) Enjoyment. These indicators reflect students' cognitive focus, emotional involvement, active participation, curiosity, and positive feelings toward English learning activities. These four indicators are measured using a 20-item Likert-scale questionnaire administered before and after the implementation of Wordwall activities.

The logical relationship between the two variables is as follows. Wordwall game activities introduce novelty, challenge, interaction, and enjoyment into vocabulary lessons, thereby creating the environmental conditions necessary to trigger students' situational interest as explained in the Four-Phase Model of Interest Development proposed by Hidi and Renninger

(2006). Through consistent and meaningful engagement in Wordwall-based activities, students are expected to develop greater attention, enthusiasm, curiosity, participation, and enjoyment in learning English. Therefore, Wordwall is assumed to have a positive effect on students' interest in learning English.

Wordwall-based activities, namely Match Up, Quiz, and Spin the Wheel, provide novelty, challenge, and immediate feedback during the learning process. According to Hidi and Renninger (2006), these learning conditions can trigger students' situational interest because they attract attention, encourage participation, and create positive learning experiences. In this study, situational interest is reflected through four observable indicators of students' interest in learning English, namely attention, enthusiasm and participation, curiosity, and enjoyment. Through consistent implementation across six instructional meetings, students' situational interest is expected to develop into maintained interest and show early indications of emerging individual interest toward learning English.

Figure 2.1 Conceptual Framework of the Relationship between Wordwall and Students' Interest in Learning English



The conceptual framework described above forms the theoretical and logical backbone of this study. The researcher hypothesizes that exposing students to Wordwall on English vocabulary will produce a significant and measurable increase in their interest in learning English, as assessed through the four indicators outlined above. It is important to note that the visual structure of the diagram in Figure 2.1 was constructed by the researcher to represent, within a single diagram, the theoretical relationship between Wordwall-based game activities (Plass et al., 2020; Adipat et al., 2021), the interactive learning environment these activities are intended to create, and the four indicators of students' interest in learning English derived from Renninger and Hidi (2016). The intellectual content, theoretical decisions, and variable relationships reflected in the diagram were determined entirely by the researcher based on the theoretical frameworks reviewed in this chapter.

D. Hypothesis

Based on the theoretical framework, the review of related literature, the review of previous studies, and the conceptual framework presented in this chapter, the hypotheses of this study are formulated as follows:

Ha (Alternative Hypothesis):

There is a significant effect of Wordwall game-based learning on students' interest in learning English, indicated by a significant difference between the pre-questionnaire and post-questionnaire mean scores ($\mu_{\text{post}} \neq \mu_{\text{pre}}$) with $p < 0.05$ (two-tailed) among Grade XI AKL students at SMK Negeri 1 Panyabungan.

H0 (Null Hypothesis):

There is no significant effect of Wordwall game-based learning on students' interest in learning English, indicated by no significant difference between the pre-questionnaire and post-questionnaire mean scores ($\mu_{\text{post}} =$

μ_{pre}) with $p > 0.05$ (two-tailed) among Grade XI AKL students at SMK Negeri 1 Panyabungan.

The alternative hypothesis (H_a) is accepted if the result of the statistical analysis shows a significance value of $p < 0.05$ (two-tailed). Conversely, the null hypothesis (H_0) is accepted if the significance value is $p > 0.05$ (two-tailed), based on the comparison between the pre-questionnaire and post-questionnaire results.



CHAPTER III

RESEARCH METHODOLOGY

A. Research Design

This study employed a quantitative approach with a pre-experimental research type. Sugiyono (2018) explains that quantitative research is a research method grounded in positivist philosophy, used to investigate specific populations or samples, in which data are collected using research instruments, analyzed statistically, and aimed at testing predetermined hypotheses. This approach was chosen because the study intends to measure the influence between two variables numerically and to test the research hypothesis empirically.

The specific design applied in this study is the One-Group Pretest-Posttest Design. According to Creswell (2014), this design involves three sequential steps: the researcher first measures the dependent variable before any treatment is introduced (pre-measurement or pre-questionnaire), then implements the treatment to the single research group across a series of instructional sessions, and finally measures the dependent variable again upon completion of the treatment (post-measurement or post-questionnaire). The difference between the two measurement results is then analyzed statistically to determine whether the treatment produced a significant effect on the dependent variable.

This design was chosen because SMK Negeri 1 Panyabungan has only one class per study program at each grade level, including the AKL program, making it impossible to form a separate control group within the same school. This condition is consistent with the characteristics of pre-experimental research, which does not involve full randomization but remains capable of demonstrating the directional effect of the treatment applied.

The pattern of this research design can be illustrated as follows:

o1 X o2

Description:

o1 = Pre-questionnaire: initial measurement of students' interest in learning English administered before the treatment

X = Treatment: implementation of Wordwall through six instructional meetings

o2 = Post-questionnaire: final measurement of students' interest in learning English administered after the treatment

The effect of Wordwall on students' interest in learning English is determined by statistically comparing the o1 (pre-questionnaire) and o2 (post-questionnaire) scores. A statistically significant increase in the post-questionnaire scores relative to the pre-questionnaire scores, confirmed through the Paired Sample t-Test or the Wilcoxon Signed-Rank Test depending on data distribution, indicates that the treatment has produced a significant positive effect on students' interest in learning English. The practical magnitude of this effect is further evaluated using Cohen's d effect size index.

Threats to Internal Validity

The researcher acknowledges that the One-Group Pretest-Posttest Design carries limitations related to internal validity that need to be explicitly addressed. This design was selected based on three practical and academic considerations. First, SMK Negeri 1 Panyabungan has only one class per study program at each grade level, including the AKL program, making random assignment into experimental and control groups within the same school unfeasible. Second, forming a comparison group from another school could introduce confounding variables such as differences in teacher quality, curriculum implementation, and school culture that are difficult to control. Third, within the scope and resources available for this undergraduate research, the One-Group Pretest-Posttest Design remains the most practical

and ethically appropriate approach for demonstrating the directional effect of the treatment.

In a one-group pretest-posttest design, several threats to internal validity must be acknowledged. Creswell (2014) identifies three major threats relevant to this type of research design, all of which were carefully considered in the present study.

The first is the history threat. According to Creswell (2014), the history threat occurs when external events that take place during the implementation period influence the dependent variable independently of the treatment, making it impossible to attribute observed changes solely to the intervention. In this study, the implementation period runs from 20 July to 12 August 2026, which falls at the very beginning of the odd semester of the 2026/2027 academic year. External events such as school orientation activities, administrative programs, national holidays, or English competitions scheduled during this period could potentially affect students' interest in learning English in ways that are unrelated to the Wordwall-based treatment. To minimize this threat, the researcher coordinated with the English teacher and the school administration prior to the research implementation to identify foreseeable scheduling conflicts, and the eight research sessions were conducted under conditions as consistent and undisturbed as possible.

The second is the maturation threat, which refers to natural changes that may occur within students over time. During the four-week implementation period, students may become more familiar with English learning activities or develop greater confidence regardless of the treatment. Therefore, the findings of this study are interpreted cautiously and are not intended to establish full causal relationships.

The third is the testing effect, which refers to the possibility that students' familiarity with the pre-questionnaire may influence their responses on the post-questionnaire beyond the effect of the treatment itself. To reduce this threat, the researcher administered the post-questionnaire

approximately five weeks after the pre-questionnaire, and employed a balanced questionnaire consisting of both positive and negative statements to minimize response bias.

These threats cannot be completely eliminated due to the single-class design of this study and are therefore acknowledged as inherent limitations of the research. Therefore, future studies are strongly encouraged to adopt a quasi-experimental design with a comparison group from a parallel class or another school in order to produce findings with stronger causal interpretability.

B. Location and Time of the Study

1. Location of the Study

This study was conducted at SMK Negeri 1 Panyabungan, located on Jalan Aek Galoga, Panyabungan District, Mandailing Natal Regency, North Sumatra Province, Indonesia. This school is one of the state vocational high schools (SMK) under the supervision of the North Sumatra Provincial Education Office and offers several vocational programs, one of which is the Accounting and Financial Institutions (AKL) program that becomes the focus of this research.

The selection of SMK Negeri 1 Panyabungan as the research location was based on several scientific and practical considerations.

First, the researcher conducted a Teaching Practicum Program (Program Pengalaman Lapangan / PPL) at SMK Negeri 1 Panyabungan from January to February 2026. During the practicum period, the researcher carried out direct classroom observation in the Grade X AKL class consisting of 27 students. The observation revealed that students' interest in learning English was considerably low. English lessons were predominantly delivered through teacher-centered methods, causing students to appear passive and to rarely participate voluntarily in classroom activities.

Second, no previous study has specifically investigated the effect of Wordwall on students' interest in learning English at SMK Negeri 1 Panyabungan. Therefore, this study has clear novelty value and provides a meaningful contribution to improving the quality of learning at the school.

Third, the researcher has sufficient access, an established relationship formed during the PPL period, and formal permission from the school to conduct this research, enabling the data collection process to be carried out systematically and ethically.

2. Time of the Study

This study was conducted across two academic years. The preliminary observation was carried out from January to February 2026, coinciding with the researcher's PPL at SMK Negeri 1 Panyabungan, during which the students were still enrolled in Grade X. The full research activities, including instrument validity and reliability testing, treatment implementation, and data collection, began on 20 July 2026, at the start of the 2026/2027 academic year. By this time, the same 27 students who participated in the preliminary observation had advanced to Grade XI. The research continues to use the same group of students as the original sample, so the sample itself remains unchanged. Only the students' grade level changes, from Grade X to Grade XI, due to the natural academic progression that occurs between the proposal seminar and the implementation of the research. The full research process, from instrument finalization to data analysis and report writing, was completed from July to August 2026.

Table 3.1. Research Schedule

No	Activity	Period
1	Preliminary observation and initial data collection (PPL)	08 January – 13 February 2026

2	Development and revision of research instrument (questionnaire)	25 April – 20 May 2026
3	Validity and reliability testing of the instrument	25 May 2026
4	Official research permission from the school	17 July 2026
5	Pre-questionnaire administration	20 July 2026
6	Treatment implementation (Wordwall game-based learning)	22 July – 10 August 2026
7	Post-questionnaire administration	12 August 2026
8	Data collection, processing, and analysis	13 – 15 August 2026
9	Final research report writing	16 August 2026

It is important to note that the schedule presented above reflects the actual timeline followed during the implementation of the research, with minor adjustments made along the way in response to field conditions, administrative processes, and other circumstances encountered during data collection.

The selection of six instructional meetings was based on several academic considerations. First, according to the academic schedule of SMK Negeri 1 Panyabungan, English is scheduled twice a week, with each meeting consisting of two 45-minute lesson periods, making eight meetings (including the pre- and post-questionnaire sessions) feasible within a four-week treatment period. Second, the vocabulary material covered in this study requires approximately six to eight meetings to ensure that students have sufficient opportunities to understand, practice, and apply the concepts through Wordwall-based activities. Third, the duration of the treatment was determined based on recommendations from the English teacher to ensure that the implementation could be integrated into the regular teaching schedule without disrupting other learning objectives.

C. Population and Sample

1. Population

Sugiyono (2018) defines population as a generalization domain consisting of objects or subjects that have certain qualities and characteristics determined by the researcher to be studied, from which conclusions are then drawn.

The population in this study consists of the same 27 students who were enrolled in Grade X AKL (Accounting and Financial Institutions) at SMK Negeri 1 Panyabungan during the 2025/2026 academic year. By July 2026, when the research was implemented, these students had advanced to Grade XI within the 2026/2027 academic year, following the normal academic progression of the school calendar. The research continues to use the same cohort of students as its population, and the sample composition remains unchanged. Only the students' grade level changes, from Grade X to Grade XI, due to the time gap between the proposal seminar and the implementation of the research.

The selection of this particular class as the research population was based on several specific considerations rather than convenience alone. First, the preliminary observation conducted during the PPL period directly identified this group of students as exhibiting low interest in learning English, characterized by passive participation and limited enthusiasm during lessons, which makes them a population for whom an interest-based intervention is genuinely needed rather than arbitrarily assigned. Second, AKL (Accounting and Financial Institutions) is a vocational program in which English functions primarily as a general subject rather than a vocationally specialized one, meaning that students' engagement with English is shaped mainly by instructional quality and classroom experience rather than by strong vocational motivation, making changes in interest more directly attributable to the treatment. Third, since SMK Negeri 1 Panyabungan operates only one class per study program at the grade level under study, this population

represents the entire available group of students with the relevant characteristics, which is consistent with the total sampling technique applied in this research.

2. Sample

Sugiyono (2018) defines a sample as a portion of the number and characteristics possessed by the population. In this study, the sampling technique used is total sampling (saturated sampling). Sugiyono (2018) explains that total sampling is applied when all members of the accessible population are used as the sample, a technique that is often practical to apply when the population size is relatively small, typically fewer than 30 individuals.

This technique was chosen primarily because the 27 students in Grade XI AKL constitute the entire accessible population of this study, that is, the whole group of students matching the population characteristics defined above, all of whom were included as the research sample rather than a portion drawn from a larger population. The fact that this population consists of only 27 students, fewer than 30, further supports the practicality of applying total sampling in this context, but is not by itself the primary reason for the choice, since the defining feature of total sampling is that the entire accessible population is used as the sample, not merely that the population happens to be small. The use of total sampling also minimizes sampling error because no member of the accessible population is excluded from the study.

It is important to acknowledge, however, that a sample of 27 students from a single class at a single school carries significant implications for the generalizability of the findings. The results of this study are context-specific and are applicable only to the population studied, namely the cohort of 27 students at SMK Negeri 1 Panyabungan who were in Grade X AKL during the proposal seminar and Grade XI AKL at the time of research implementation in the 2026/2027 academic year. The findings cannot be generalized broadly to other SMK classes

or schools in Indonesia without further replication studies involving larger and more diverse samples. This limitation is acknowledged as an inherent constraint of the total sampling approach in a single-class research context.

Table 3.2. Research Sample Data

No	Class	Study Program	Number of Students	Note
1	XI AKL	Accounting and Financial Institutions	27 students	Research sample (total sampling)
		Total	27 students	

The findings of this study are applicable only to the specific cohort of students who participated in the research, namely the same 27 students who were in Grade X AKL during the proposal seminar and who advanced to Grade XI AKL at SMK Negeri 1 Panyabungan by the time of research implementation in the 2026/2027 academic year. Therefore, the results cannot be generalized to other schools, study programs, or grade levels without further replication studies involving different samples and research contexts.

D. Technique of Data Collection

1. Instrument of Data Collection

The data collection instrument used in this study is a closed-ended questionnaire on students' interest in learning English. Sugiyono (2018) explains that a questionnaire is a data collection technique conducted by providing a set of written questions or statements to respondents to be answered. This technique is efficient when the researcher clearly knows the variables to be measured.

The questionnaire in this study was developed based on four indicators of students' interest in learning English derived from the

theoretical framework of Hidi and Renninger (2006), namely: (1) Attention, (2) Enthusiasm and Participation, (3) Curiosity, and (4) Enjoyment. These four indicators were selected because they comprehensively represent the affective and behavioral dimensions of students' interest that are relevant and observable in a Wordwall environment. They are also aligned with the learning experiences produced through Match Up, Quiz, and Spin the Wheel activities on vocabulary material implemented across six instructional meetings at SMK Negeri 1 Panyabungan.

The questionnaire uses a four-point Likert Scale. The decision to use four response options instead of five was made deliberately to avoid students' tendency to select the neutral midpoint, which does not provide meaningful directional information for research purposes. The scoring for each response option is presented as follows:

Table 3.3. Likert Scale Scoring

Response Option	Code	Score for Positive Statement (+)	Score for Negative Statement (-)
Strongly Agree	SA	4	1
Agree	A	3	2
Disagree	D	2	3
Strongly Disagree	SD	1	4

This questionnaire was administered to all 27 students of Grade XI AKL in two stages: before the treatment (pre-questionnaire) to identify students' initial level of interest, and after the treatment (post-questionnaire) to measure changes in students' interest following the implementation of Wordwall across six instructional meetings.

2. Instrument

The following table presents the sub-indicators of students' interest in learning English questionnaire used in this study. The

instrument consists of 20 items distributed across four indicators, with a balanced distribution of 10 positive (+) and 10 negative (-) statements to minimize acquiescence bias, which refers to the tendency of respondents to agree with statements regardless of their content (Podsakoff et al., 2003).

Table 3.4. Sub-Indicators of Students' Interest Questionnaire
Grade XI AKL, SMK Negeri 1 Panyabungan

Total: 20 items | Distribution: 10 positive (+) statements, 10 negative (-) statements

No	Indicator	Sub-Indicator	Positive Item Numbers	Negative Item Numbers	Total
1	Attention	Focusing and maintaining concentration during Wordwall-based English lessons	1, 2	3	3
		Not being easily distracted during English learning activities	4	5	2
	Subtotal		3	2	5
2	Enthusiasm and Participation	Responding actively and enthusiastically to Wordwall game activities	6	7, 8	3
		Participating voluntarily in English	9	10	2

		learning tasks without being prompted			
	Subtotal		2	3	5
3	Curiosity	Seeking additional information about vocabulary beyond game sessions	11, 12	13	3
		Voluntary question-asking about vocabulary encountered in Wordwall activities	14	15	2
	Subtotal		3	2	5
4	Enjoyment	Feeling happy and satisfied during English lessons using Wordwall	16	17, 18	3
		Viewing Wordwall-based English lessons as a pleasant and rewarding experience	19	20	2
	Subtotal		2	3	5
	Grand Total		10 items	10 items	20

Source: Adapted from (Renninger and Hidi 2016)

The distribution of positive and negative items across the four indicators was designed to support balanced measurement within each sub-indicator of students' interest in learning English.

The use of both positive and negative statements aims to minimize acquiescence bias, which refers to the tendency of respondents to agree with statements regardless of their actual content (Podsakoff et al., 2003). By balancing the statements, students are encouraged to read each item carefully, thereby improving the reliability and accuracy of the questionnaire.

Furthermore, the distribution of items was adjusted to the characteristics of each indicator and sub-indicator. Indicators related to positive emotional responses, such as Attention and Enjoyment, contain more positive statements, while indicators associated with participation and curiosity, namely Enthusiasm and Participation and Curiosity, include more negative statements to identify the lack of active engagement more effectively. Overall, the questionnaire maintains a balanced total of 10 positive and 10 negative items to ensure a more comprehensive and objective measurement of students' interest in learning English.

The questionnaire used in this study was developed based on the indicators of students' interest proposed by Hidi and Renninger (2006) and Renninger and Hidi (2016). Prior to its use in the study, the instrument underwent an expert validation process to ensure the appropriateness of its content, clarity of language, and suitability of each item in measuring students' interest in learning English. The validation sheet and the results of the expert judgment are provided in the appendix of this thesis.

In addition to expert validation, the instrument underwent statistical validity and reliability testing through a try-out prior to its use in the main study. The validity test was conducted using the Pearson Product Moment correlation technique, while the reliability test was carried out using Cronbach's Alpha coefficient with the assistance of IBM SPSS Statistics. These procedures were intended to ensure that the instrument possessed adequate validity and

reliability for measuring students' interest in learning English. The results of the try-out are presented in the following subsections.

3. Validity Test

Before the questionnaire was used in the actual study, each item was subjected to a validity test to ensure it truly measures what it is intended to measure. Sugiyono (2018) states that a valid instrument means the measuring tool can be used to measure what it is supposed to measure.

The validity test used the **Pearson Product Moment Correlation** technique with the assistance of IBM SPSS Statistics. If $r\text{-count} > r\text{-table}$ at the 5% significance level, the item is declared valid. If $r\text{-count}$ is less than or equal to $r\text{-table}$, the item is declared invalid and must be revised or removed from the instrument.

Based on the try-out of the 20-item questionnaire involving 30 respondents, the corrected item-total correlation ($r\text{-count}$) for all 20 items ranged from .416 to .722, exceeding the $r\text{-table}$ value of .361 at the 5% significance level ($df = 28$). Therefore, all 20 items were declared valid and retained to measure students' interest in the pre-questionnaire and post-questionnaire administration.

The complete item-by-item validity results, listing the corrected item-total correlation ($r\text{-count}$), the $r\text{-table}$ value, and the validity decision for each of the 20 items individually, are presented in Appendix 3 (SPSS Item-Total Statistics Output) rather than being reproduced item-by-item in the body text, so that the discussion here remains focused while still allowing readers to verify the validity of every single item.

4. Reliability Test

The reliability test was conducted to determine the extent to which the instrument is dependable and produces consistent results when used with the same subjects under different conditions. Sugiyono (2018) states that a reliable instrument produces the same data when used to measure the same object multiple times.

The reliability test used the **Cronbach's Alpha** method. An instrument is declared reliable if its Cronbach's Alpha value is equal to or greater than 0.60.

Table 3.5. Criteria for Interpreting Reliability Coefficients

Coefficient Interval	Reliability Level
0.00 - 0.199	Very Low
0.20 - 0.399	Low
0.40 - 0.599	Moderate
0.60 - 0.799	Strong
0.80 - 1.000	Very Strong

Source: Sugiyono (2018)

The reliability test of the 20 valid items yielded a Cronbach's Alpha value of 0.903. Referring to the criteria in Table 3.5, this value falls within the Very Strong category, indicating that the questionnaire is highly reliable for measuring students' interest in learning English.

It is important to clarify that this Cronbach's Alpha value of 0.903 represents the reliability of the questionnaire as a whole (all 20 items combined), not the reliability of each of the four indicator subscales (Attention, Enthusiasm and Participation, Curiosity, Enjoyment) individually. Because each subscale contains only 4-6 items, a formal per-subscale reliability analysis was not conducted for the try-out data; however, since the overall instrument demonstrates a Very Strong level of internal consistency and the sub-indicator structure in Table 3.4 is directly derived from the theoretical framework of Renninger and Hidi (2016), the researcher considers the overall reliability coefficient sufficient to support the use of the four subscales for the descriptive and inferential indicator-level analysis reported in Chapter IV. Future studies using this instrument are nonetheless encouraged to report subscale-level reliability where the sample size allows for a more stable estimate.

5. Data Collection Procedure

Data collection was carried out through three main stages:

a. **Stage One: Pre-Questionnaire**

Before the treatment was implemented, the researcher distributed a students' interest questionnaire to 27 students of Grade XI AKL at SMK Negeri 1 Panyabungan. The pre-questionnaire aimed to identify the initial condition of students' interest in English learning before the implementation of Wordwall. Students were given sufficient time to complete the questionnaire honestly, independently, and without any pressure from any party.

b. **Stage Two: Treatment.**

The researcher implemented Wordwall across six instructional meetings. Full details of the treatment plan are presented in Sub-section 6 below.

The implementation of game-based learning through Wordwall in this study was adapted to a technical constraint at the school, namely the limited availability of projectors (infocus). Since SMK Negeri 1 Panyabungan has only five projector units shared across its 27 classrooms, and these units were also needed by other classes and subjects at the same hours, the researcher could only secure projector access for two of the six treatment meetings through advance scheduling with the school's facility coordinator. In these two meetings, the Wordwall games were displayed on a shared screen at the front of the class so that all students could follow the activity collectively. Throughout all six meetings, the students' regular English teacher remained the instructor of record and continued to lead the core classroom instruction, so as not to disrupt students' usual learning routine with an unfamiliar teaching figure; the researcher's role was limited to administering the Wordwall treatment segment of each lesson under the teacher's supervision.

In the remaining four meetings, without projector access, students accessed and completed the Wordwall activities directly through their personal smartphones, guided by the researcher's verbal instructions for that session's activity.

Regardless of whether the projector was available, students participated in the learning process in small groups, which allowed them to engage in discussion, collaboration, and interaction with their group members in completing the tasks provided through Wordwall activities, either around the shared screen or around a group member's smartphone.

A stable internet connection was required throughout the six meetings to ensure smooth access to the Wordwall platform, whether through the shared screen or through students' personal smartphones. In addition, attendance varied slightly across the six meetings, as a small number of students were occasionally absent or unwell on a given meeting day, meaning that not every student was present for all six treatment sessions. This attendance variation and the limited projector availability are discussed further as constraints of this study in Chapter IV.

This arrangement was designed to create an interactive and collaborative learning environment while also adjusting to the available classroom facilities.

c. Stage Three: Post-Questionnaire .

After all treatment sessions were completed, the same questionnaire was distributed again to the same 27 students. The post-questionnaire aimed to measure students' interest in English learning after the implementation of Wordwall. The pre-questionnaire and post-questionnaire scores were then compared and analyzed statistically to test the research hypothesis.

6. Treatment Plan

The study was conducted across eight research meetings over a four-week period, from 20 July to 12 August 2026, consisting of one meeting for the administration of the pre-questionnaire on 20 July 2026, six instructional meetings for the implementation of Wordwall from 22 July to 10 August 2026, and one meeting for the administration of the post-questionnaire on 12 August 2026. Each meeting lasted 90 minutes (2×45 minutes). The material covered was narrative text, comprising reading, listening, text structure and language features, and grammar (the simple past tense), based on the Grade XI Erlangga SPLASH English textbook for the Kurikulum Merdeka, which is part of the English curriculum for the odd semester of the 2026/2027 academic year. Wordwall activities were used throughout to reinforce the key vocabulary embedded in each of these materials.

The treatment was designed progressively around two narrative texts. Meetings 1 and 2 introduced the narrative text The Legend of Malin Kundang through reading and listening activities, meetings 3 and 4 introduced the narrative text Rawa Pening through reading and further narrative comprehension practice, and meetings 5 and 6 focused on the structure, language features, and grammar (the simple past tense) drawn from both narrative texts. Wordwall activities were used in every meeting to reinforce the vocabulary relevant to that session's material.

Table 3.6. Wordwall Treatment Plan

Meeting	Day	Topic	Wordwall Activity	Learning Objective
1	Wednesday, 22 July 2026	Narrative Text: The Legend of Malin Kundang	Quiz and Match Up	Students can understand the main ideas and important vocabulary in the narrative text The Legend of Malin Kundang

				and identify the meaning of words through Wordwall activities.
2	Monday, 27 July 2026	Listening Section: The Legend of Malin Kundang – based on Activities 1 and 2, page 3, from the Erlangga textbook	Spin the Wheel	Students can listen to and understand information from the narrative text The Legend of Malin Kundang and respond appropriately to questions based on the listening activity.
3	Wednesday, 29 July 2026	Reading Section: Rawa Pening, based on Activity 6 from the Erlangga textbook	Spin the Wheel	Students can understand the main ideas, specific information, and vocabulary in the narrative text Rawa Pening through reading activities.
4	Monday, 03 August 2026	Narrative Text: Rawa Pening	Quiz	Students can identify the main ideas, characters, events, and important information in the narrative text Rawa Pening through Wordwall quiz activities.
5	Wednesday, 05 August 2026	Structure and Language Features of Narrative Text	Spin the Wheel	Students can identify and understand the generic structure and language features of narrative texts,

				including orientation, complication, resolution, and the use of past tense and action verbs.
6	Monday, 10 August 2026	Simple Past Tense: Second Form (V2) in the Story of Malin Kundang	Quiz	Students can identify and use the second form of verbs (V2) found in the narrative text The Legend of Malin Kundang through Wordwall quiz activities.

To provide a clearer description of the treatment activities, examples of the Wordwall tasks used in this study are presented as follows. In the Match Up activity, students were asked to match target vocabulary words with their correct meanings or definitions, for example matching the word “curious” with “having a strong desire to know something” and “enormous” with “very large in size.” In the Quiz activity, students answered multiple-choice questions about vocabulary meaning and usage, such as: “What does the word ‘monotonous’ mean? A) exciting B) repetitive and boring C) loud,” with the correct answer being “repetitive and boring.” In the Spin the Wheel activity, students responded to randomly selected vocabulary words and were asked to define them or use them in a sentence. For example, when the wheel stopped at the word “participate,” students were expected to provide its meaning or construct a sentence using the word. These examples illustrate how Wordwall activities were used to reinforce students’ vocabulary understanding in an interactive and engaging manner. Each meeting followed a consistent three-phase structure. The warm-up phase (10 minutes) involved the teacher reviewing the previous lesson and

introducing the Wordwall activity for the day, ensuring that students understood the rules and objectives of the game.

The game activity phase (65 minutes) involved students actively participating in the Wordwall activities displayed on the classroom screen, while the teacher acted as a facilitator by monitoring participation and providing guidance when needed. Within this 65-minute phase, each meeting used exactly one Wordwall activity set corresponding to that session's designated activity type (Match Up, Quiz, or Spin the Wheel, as specified in Table 3.6), so that the format and structure of the game activity were kept consistent across the six meetings; the specific number of vocabulary rounds per set followed the content available in that session's textbook material.

The reflection and feedback phase (15 minutes) involved the teacher reviewing the game results displayed on Wordwall, discussing common errors in identifying vocabulary, and explicitly linking the game activities to that session's specific learning objective as presented in Table 3.6, whether related to reading, listening, text structure, or grammar.

This consistent three-phase structure was applied across all six meetings to create a structured and stable learning environment, minimizing confounding variables that could influence students' interest measurement, while maintaining engagement and novelty through game-based learning.

E. Technique of Data Analysis

The data analysis in this study encompasses four stages: descriptive statistical analysis, prerequisite normality test, hypothesis testing using the Paired Sample t-Test, and effect size calculation using Cohen's d. All processes were conducted with IBM SPSS Statistics version 23 or newer.

1. Descriptive Statistical Analysis

Descriptive statistical analysis was used to describe the characteristics of the pre-questionnaire and post-questionnaire data collected from 27 students. According to Sugiyono (2018), descriptive statistics analyze data by describing or presenting the collected data as they are, without drawing broader generalizations.

The analysis includes the calculation of mean, minimum value, maximum value, and standard deviation of both pre-questionnaire and post-questionnaire scores. Student interest levels are classified using the following categories:

The score interval categories in Table 3.7 were determined using the equal interval formula proposed by Sudjana (2016):

$$I = \frac{R}{K}$$

Where:

I = interval width

R = range (maximum score – minimum score)

K = number of categories

In this study, the maximum possible score was 80 and the minimum possible score was 20, therefore:

$$R = 80 - 20 = 60$$

The number of categories used in this study is 3, so:

$$I = \frac{60}{3} = 20$$

Thus, each category interval has a width of 20 points.

Table 3.7. Categories of Students' Interest in Learning English

Score Interval	Category
61 – 80	High
41 – 60	Moderate
20 – 40	Low

Note: Based on 20 items, maximum score = 80, minimum score = 20.

The score intervals presented in Table 3.7 were determined using the equal interval formula proposed by Sudjana (2016). The calculation is based on a questionnaire consisting of 20 items with a 4-point Likert scale, which yields a minimum possible score of 20 and a maximum possible score of 80, producing a total range of 60 points. Since three categories are used in this study, the interval width for each category is calculated as 60 divided by 3, which equals 20 points per category. These three categories, namely High (61-80), Moderate (41-60), and Low (20-40), were adopted because three broad categories are sufficient to differentiate students' levels of interest in a meaningful and interpretable way without creating artificial distinctions within a small sample of 27 students.

2. Prerequisite Test: Normality Test

Before hypothesis testing, the data were first tested for normality to ensure that the difference scores between the post-questionnaire and pre-questionnaire results (post-pre) were normally distributed, allowing the appropriate statistical test to be determined. Since the sample size in this study was 27 students ($n < 50$), the Shapiro-Wilk test was used, as it is considered more accurate and sensitive than the Kolmogorov-Smirnov test for small sample sizes.

If the normality test results show a significance value (Sig.) greater than 0.05, the data are considered normally distributed, and the hypothesis is tested using the Paired Sample t-Test. However, if the significance value (Sig.) is less than 0.05, indicating that the data are not

normally distributed, the Wilcoxon Signed-Rank Test is used as a non-parametric alternative. The Wilcoxon Signed-Rank Test is appropriate for comparing two related measurements when the assumption of normality is not fulfilled.

Decision criteria:

- If Sig. > 0.05: the data are normally distributed → proceed with the Paired Sample t-Test.
- If Sig. < 0.05: the data are not normally distributed → use the Wilcoxon Signed-Rank Test as a non-parametric alternative.

3. Hypothesis Testing: Paired Sample t-Test

Hypothesis testing in this study used the Paired Sample t-Test. This test compares the mean scores of two measurements from the same subjects, namely the pre-questionnaire and post-questionnaire scores of the 27 Grade XI AKL students, to determine whether there is a statistically significant difference in students' interest in learning English before and after the implementation of Wordwall.

Before conducting the Paired Sample t-Test, the normality assumption was examined using the difference scores between the post-questionnaire and pre-questionnaire results (post-pre). If the difference scores are normally distributed (Sig. > 0.05), the Paired Sample t-Test is used. However, if the difference scores are not normally distributed (Sig. < 0.05), the Wilcoxon Signed-Rank Test is employed as a non-parametric alternative.

The formula for the Paired Sample t-Test is as follows:

$$t = \frac{\bar{D}}{SD/\sqrt{n}}$$

Description:

- t = calculated t-value

- $\bar{D}$ = mean of the differences between post-questionnaire and pre-questionnaire scores
- SD = standard deviation of the difference scores
- n = number of samples (27 students)

Decision criteria:

- If Sig. (2-tailed) < 0.05: H_0 is rejected and H_a is accepted. This indicates a significant effect of Wordwall on students' interest in learning English at Grade XI AKL, SMK Negeri 1 Panyabungan.
- If Sig. (2-tailed) > 0.05: H_0 is accepted and H_a is rejected. This indicates no significant effect of Wordwall on students' interest in learning English at Grade XI AKL, SMK Negeri 1 Panyabungan.

4. Effect Size Calculation: Cohen's d

To measure the practical magnitude of the effect, this study used Cohen's d as the effect size index. Cohen's d is the most appropriate and widely used measure for paired-sample designs involving Likert-scale affective instruments (Cohen, 1988; Field, 2018).

Formula:

$$d = \frac{M_{post} - M_{pre}}{SD_{diff}}$$

Description:

- d = Cohen's d effect size value
- M_{post} = mean post-questionnaire score
- M_{pre} = mean pre-questionnaire score
- SD_{diff} = standard deviation of the difference scores (post – pre)

Table 3.8. Interpretation Criteria for Cohen's d

Cohen's d Value	Effect Size Category
$d < 0.20$	Negligible
$0.20 \leq d < 0.50$	Small
$0.50 \leq d < 0.80$	Medium
$d \geq 0.80$	Large

Source: Cohen (1988)

In the context of this study, the interpretation of Cohen's d is not only statistical but also practical. A Cohen's d value of 0.80 or higher (large effect) indicates a substantial improvement in students' interest in learning English, which may be reflected in a shift from a lower interest category to a higher one, for example from Low to High. A Cohen's d value between 0.50 and 0.80 (medium effect) indicates a noticeable improvement that can be observed in classroom participation, attention, enthusiasm, and enjoyment during English learning activities. Meanwhile, a Cohen's d value below 0.20 (negligible effect) indicates that the treatment has very little practical impact on students' interest in learning English.

A Cohen's d value of 0.80 or higher indicates a large and practically significant effect of Wordwall on students' interest in learning English. This not only confirms statistical significance but also demonstrates practical significance in real classroom practice. By combining the results of the Paired Sample t-Test and Cohen's d, this study provides a comprehensive explanation of both statistical significance and practical significance of the intervention.

5. Effect Size for the Wilcoxon Signed-Rank Test: r

The Cohen's d formula above is used when the Paired Sample t-Test is selected, that is, when the pre-questionnaire and post-questionnaire difference scores are normally distributed. If the normality assumption is not met and the Wilcoxon Signed-Rank Test is used

instead, Cohen's d is no longer the statistically appropriate effect size index, since it is a parametric measure derived from means and standard deviations (Cohen, 1988; Field, 2018). In that case, this study uses the rank-based correlation coefficient r as the effect size index, calculated by dividing the Z statistic obtained from the Wilcoxon test by the square root of the number of paired observations (Field, 2018).

Formula:

$$r = Z / \sqrt{N}$$

Description:

r = effect size

Z = the Wilcoxon Signed-Rank Test statistic

N = number of paired observations

Table 3.9. Interpretation Criteria for Effect Size r

r Value	Effect Size Category
$r < 0.10$	Negligible
$0.10 \leq r < 0.30$	Small
$0.30 \leq r < 0.50$	Medium
$r \geq 0.50$	Large

Source: Cohen (1988)

In this study, the choice between Cohen's d and r depends entirely on the outcome of the normality test reported in Chapter IV. Whichever index applies, it is interpreted using the same general logic: a larger value indicates a greater practical magnitude of the effect of Wordwall on students' interest in learning English, independent of the statistical significance result.

CHAPTER IV

RESEARCH FINDINGS AND DISCUSSION

A. Description of the Research Setting

This research was conducted at SMK Negeri 1 Panyabungan, a public vocational high school located on Jalan Sukaramai Aek Galoga, Pidoli Lombang Village, Panyabungan Kota District, Mandailing Natal Regency, North Sumatra Province. The research involved 27 students of Grade XI AKL (Accounting and Financial Institutions) during the odd semester of the 2026/2027 academic year. As explained in Chapter III, this class was the only AKL class at the Grade XI level in the school, so all 27 students were included through total sampling; Table 4.1 shows how this cohort fits within the school's overall grade-level enrollment.

Table 4.1 Number of Students by Grade

Grade	Number of Students
Grade X	232
Grade XI	187
Grade XII	147
Total	566

Source: SMK Negeri 1 Panyabungan school data (2026).

The participants in this research were the 27 students of Grade XI AKL, one of the classes under the Grade XI cohort shown in Table 4.1. Regarding facilities relevant to the implementation of the treatment, Table 4.2 summarizes the school's main facilities, most notably its shared projector units.

Table 4.2 Main Facilities of SMK Negeri 1 Panyabungan

No.	Facility	Quantity	Condition
1	Classrooms	27	Good

2	Computer Laboratories	3	Good
3	Computers	100	Good
4	Smart TV	1	Good
5	Projectors	5	Good
6	Whiteboards	48	Good
7	Library	1	Good
8	Teachers' Room	1	Good
9	Principal's Room	1	Good
10	Administration Room	1	Good
11	Counseling Room	1	Good
12	Wi-Fi	1	Good
13	Hall	1	Good
14	Mosque	1	Good
15	Toilets	11	Good

Source: SMK Negeri 1 Panyabungan school data (2026).

As shown in Table 4.2, the school has five projector (infocus) units shared across its 27 classrooms. Because these units were also needed by other classes at the same hours, the researcher was only able to secure projector access for two of the six treatment meetings, during which the Wordwall activities were displayed on a shared screen; in the remaining four meetings, students accessed the Wordwall activities directly through their personal smartphones connected to the school's Wi-Fi network. This constraint, together with minor attendance variation across meetings, is discussed further in Section F of this chapter.

The full research process was carried out across eight research meetings, consisting of one meeting for the administration of the pre-questionnaire, six instructional meetings in which the treatment was implemented using Wordwall activities (Match Up, Quiz, and Spin the Wheel), and one meeting for the administration of the post-questionnaire, following the schedule described in Chapter III.

B. Description of the Data

This section presents the descriptive statistics of students' interest in learning English before and after the implementation of Wordwall, based on the pre-questionnaire and post-questionnaire scores of 27 Grade XI AKL students, processed using IBM SPSS Statistics.

1. Descriptive Statistics of Pre-Questionnaire and Post-Questionnaire Scores

Table 4.3 Descriptive Statistics of Pre-Questionnaire and Post-Questionnaire Scores

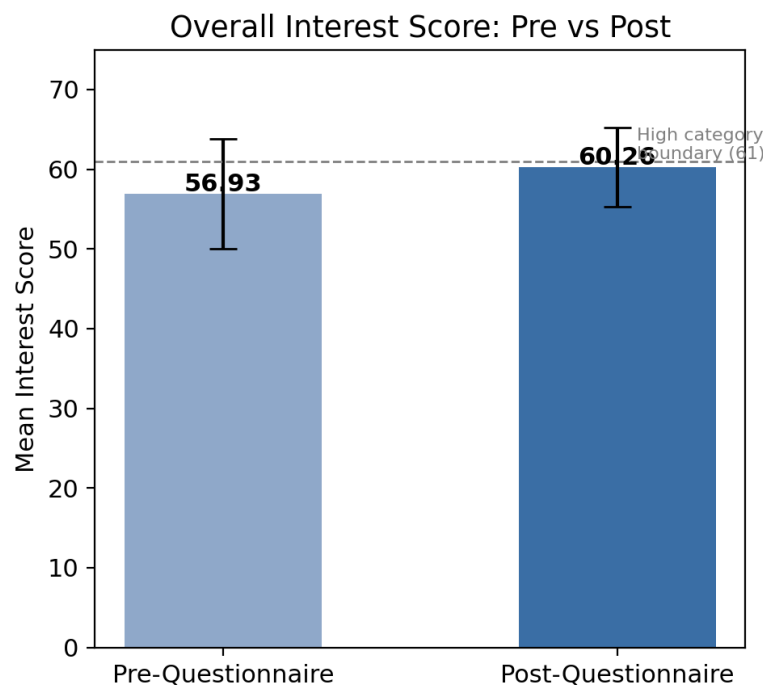
Score	N	Minimum	Maximum	Mean	Std. Deviation
Pre-Questionnaire	27	48.00	73.00	56.93	6.878
Post-Questionnaire	27	45.00	70.00	60.26	4.966

Source: IBM SPSS Statistics output, processed by the researcher (2026).

As shown in Table 4.3, the mean score of students' interest increased from 56.93 in the pre-questionnaire to 60.26 in the post-questionnaire, an increase of 3.33 points. According to the score interval categories established in Table 3.7 of Chapter III (High 61-80, Moderate 41-60, Low 20-40), both the pre-questionnaire and post-questionnaire mean scores fall within the Moderate category, although the post-questionnaire mean is close to the boundary of the High category. It is worth noting that the post-questionnaire minimum score (45) is lower than the pre-questionnaire

minimum score (48), even though the group mean increased; this is not a contradiction, since a mean-level increase describes the sample as a whole and does not guarantee that every individual student's score rose, and this single lowest post-questionnaire score belongs to one of the eight students discussed later in Section F.1 whose scores declined rather than reflecting a general downward trend. The standard deviation decreased from 6.878 in the pre-questionnaire to 4.966 in the post-questionnaire, indicating that students' scores became more homogeneous after the treatment. This homogenization worked from both directions: consistent with the narrower minimum-to-maximum range in the post-questionnaire (45–70) compared to the pre-questionnaire (48–73), the treatment appears to have raised the interest of students who started lowest while slightly moderating the scores of students who started highest, rather than simply shifting the whole distribution upward by an equal amount.

Figure 4.1 Overall Interest Score, Pre- and Post-Questionnaire



2. Category Distribution of Students' Interest

To provide a more detailed picture of how individual students' interest scores were distributed across the three categories established in Table 3.7, the raw score distribution recorded in the SPSS stem-and-leaf output was classified into the High, Moderate, and Low categories. The results are presented in Table 4.4.

Table 4.4 Category Distribution of Students' Interest in the Pre-Questionnaire and Post-Questionnaire

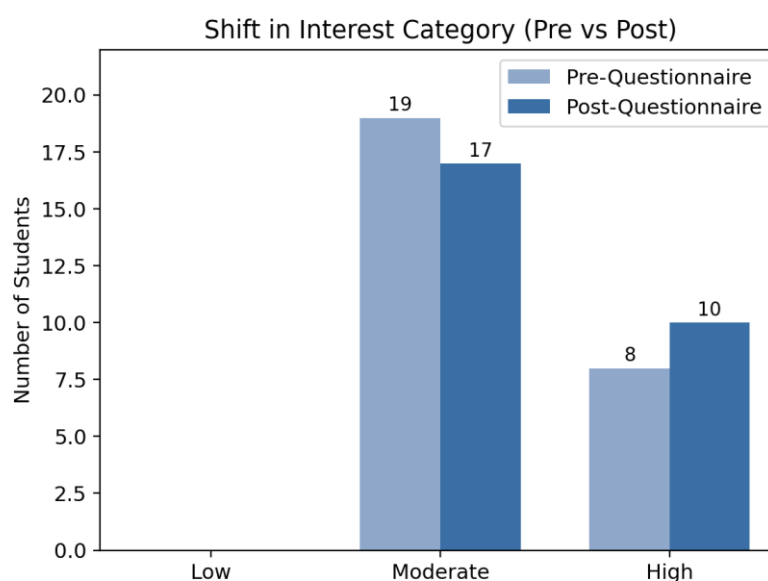
Category	Score Interval	Pre-test (n)	Pre-test (%)	Post-test (n)	Post-test (%)
High	61-80	8	29.6%	10	37.0%
Moderate	41-60	19	70.4%	17	63.0%
Low	20-40	0	0.0%	0	0.0%
Total	-	27	100.0%	27	100.0%

Source: processed by the researcher based on the SPSS descriptive output (2026).

Table 4.4 shows that no student fell into the Low category in either measurement, which suggests that students entered the treatment with at least a moderate level of interest in English. The number of students in the High category increased from 8 students (29.6%) in the pre-questionnaire to 10 students (37.0%) in the post-questionnaire, while the number of students in the Moderate category decreased from 19 (70.4%) to 17 (63.0%). This shift is consistent with the increase in the mean score reported in Table 4.3 and indicates that a portion of the sample moved toward a higher level of interest following the Wordwall-based treatment, although the majority of students remained within the Moderate category. The complete absence of students in the Low category, even before the treatment, is worth examining alongside the preliminary observation reported in Chapter I, which described students during the Teaching Practicum Program (PPL) in

January–February 2026 as largely passive and showing limited enthusiasm at the start of English lessons, a pattern the researcher interpreted at the time as indicating low interest. Taken together, this points to a plausible gap between qualitative classroom observation and quantitative self-report: a Likert-scale questionnaire measuring general statements about attention, enthusiasm, curiosity, and enjoyment may not fully capture moment-to-moment classroom passivity of the kind noted during PPL, and students completing a self-report instrument may also be inclined to select moderate rather than extreme responses, a general tendency in Likert-scale research sometimes referred to as central tendency bias (Nadler, Weston, & Voyles, 2015). It is also possible that the PPL observation, conducted in a different class (Grade X AKL) and several months before this study's own pre-questionnaire (Grade XI AKL), captured a genuinely different group of students at a different point in time, so the two data sources are not directly comparable on a one-to-one basis. Either way, this discrepancy suggests that future studies would benefit from pairing a Likert-scale interest questionnaire with a classroom observation checklist administered concurrently, so that self-reported and observed interest can be compared within the same sample and timeframe.

Figure 4.2 Shift in Interest Category, Pre- and Post-Questionnaire



C. Prerequisite Test: Normality Test

Prior to hypothesis testing, a normality test was conducted on the difference scores between the post-questionnaire and pre-questionnaire results (post-pre) using the Shapiro-Wilk test, in accordance with the procedure established in Chapter III, since the sample size of this study ($n = 27$) is below 50.

Table 4.5 Results of the Shapiro-Wilk Normality Test

Score	Statistic	df	Sig.
Difference (Post-Pre)	0.986	27	0.967

Source: IBM SPSS Statistics output, processed by the researcher (2026).

Based on the decision criteria established in Chapter III, the difference scores are considered normally distributed if the significance value (Sig.) is greater than 0.05. Table 4.5 shows that the difference scores (post-questionnaire minus pre-questionnaire) obtained a Shapiro-Wilk significance value of 0.967, which is well above 0.05, indicating that the difference scores are normally distributed. It is worth noting that this differs from testing the pre-questionnaire and post-questionnaire scores separately, which yielded borderline significance values of 0.050 and 0.048; however, in accordance with the procedure established in Chapter III, the relevant normality assumption for a Paired Sample t-Test concerns the distribution of the difference scores, not the two sets of raw scores individually. Since the normality assumption was satisfied, this study proceeded with the parametric Paired Sample t-Test, consistent with the decision rule established in Chapter III.

D. Hypothesis Testing: Paired Sample t-Test

Because the normality assumption of the difference scores was satisfied, the hypothesis of this study was tested using the Paired Sample t-Test, a parametric statistical test appropriate for comparing two related

measurements when the difference scores are normally distributed (Field, 2018). This test compares the pre-questionnaire and post-questionnaire scores of the 27 Grade XI AKL students to determine whether Wordwall had a statistically significant effect on students' interest in learning English.

Table 4.6 Paired Samples Statistics

	Mean	N	Std. Deviation	Std. Error Mean
Pre-Questionnaire	56.93	27	6.878	1.324
Post-Questionnaire	60.26	27	4.966	0.956

Source: IBM SPSS Statistics output, processed by the researcher (2026).

Table 4.6 shows the descriptive statistics for the pre-questionnaire and post-questionnaire scores of the 27 students used as the basis for the Paired Sample t-Test. The post-questionnaire mean (60.26) is higher than the pre-questionnaire mean (56.93), consistent with the descriptive results reported in Table 4.3.

Table 4.7 Paired Samples Test (Post-Questionnaire – Pre-Questionnaire)

Mean Difference	Std. Deviation	Std. Error Mean	95% CI Lower	95% CI Upper	t	df	Sig. (2-tailed)
3.333	7.206	1.387	0.483	6.184	2.404	26	0.024

Source: IBM SPSS Statistics output, processed by the researcher (2026).

Table 4.7 shows a mean difference of 3.333 (post-questionnaire minus pre-questionnaire), with a standard deviation of 7.206 and a standard error of the mean of 1.387. The 95% confidence interval for the mean difference ranges from 0.483 to 6.184, which does not include zero, indicating a genuine increase in scores. The t-test produced a t value of 2.404 with 26 degrees of freedom and a Sig. (2-tailed) value of 0.024. Referring to the decision criteria established in Chapter III, since the significance value (0.024) is smaller than 0.05, the null hypothesis (H₀) is

rejected and the alternative hypothesis (H_a) is accepted. This result indicates that there is a statistically significant increase in students' interest in learning English at Grade XI AKL, SMK Negeri 1 Panyabungan following the implementation of Wordwall, with post-questionnaire scores significantly higher than pre-questionnaire scores.

E. Effect Size

Chapter III proposed Cohen's d as the effect size index for this study, to be used if the data satisfy the normality requirement associated with the Paired Sample t -Test. Since the normality test reported in Table 4.5 showed that the difference scores were normally distributed (Sig. = 0.967), Cohen's d is the statistically appropriate effect size index for the present data, as it is a parametric measure derived from the mean and standard deviation of the difference scores under the assumption of normality (Cohen, 1988; Field, 2018).

The formula used is as follows:

$$d = \bar{D} / SD$$

Description:

d = Cohen's d effect size value

$\bar{D}$ = mean of the difference scores (post-questionnaire minus pre-questionnaire)

SD = standard deviation of the difference scores

Based on the mean difference and standard deviation of the difference scores reported in Table 4.7, the effect size is calculated as follows:

$$d = 3.333 / 7.206 = 0.463$$

The value of Cohen's d (0.463) was then interpreted using the criteria in Table 3.8. A Cohen's d value of 0.463 falls within the Small category ($0.20 \leq d < 0.50$), close to the boundary of the Medium category ($0.50 \leq d < 0.80$).

This result indicates that beyond being statistically significant, the effect of Wordwall on students' interest in learning English is of a small-to-

moderate practical magnitude. In the context of classroom practice, an effect size close to the medium boundary suggests a change that may be noticeable to teachers, reflected in the increase in mean score and the movement of several students from the Moderate to the High interest category shown in Table 4.4, although the magnitude of the change is more modest than a fully medium or large effect would indicate. This magnitude can also be expressed in terms of variance explained. Using the standard conversion of Cohen's d to the point-biserial correlation coefficient, $r = d / \sqrt{(d^2 + 4)}$ (Cohen, 1988; Rosenthal, 1994), a d of 0.463 corresponds to $r \approx 0.225$, so $r^2 \approx 0.051$, meaning that roughly 5% of the variance in students' change in interest scores is associated with the Wordwall treatment as measured in this study. The remaining variance is attributable to other factors not measured here, such as individual differences in prior interest, personality, or classroom conditions on a given day. A treatment explaining around 5% of variance in an affective outcome after only six meetings, without a control group to net out non-treatment influences such as maturation or general novelty, is consistent with what Cohen (1988) considers a small-to-moderate effect in behavioral research, where the majority of variance in complex, multiply-determined constructs such as interest is rarely captured by a single classroom intervention.

F. Discussion

Before turning to the research questions, it is important to note upfront that this Discussion is read against two practical constraints in how the treatment was delivered: projector access was shared with other classes and therefore unavailable for every meeting, and the researcher's role during each session was limited to delivering the Wordwall treatment, while the classroom English teacher retained instructional control throughout so that the regular flow of the lesson was not disrupted. These constraints, detailed fully in Section F.6, are referenced at several points below where they help

explain a specific finding, rather than being treated only as a closing disclaimer.

1. The Effect of Wordwall on Students' Interest

The first research question of this study asked whether there is a significant effect of Wordwall on students' interest in learning English at Grade XI AKL, SMK Negeri 1 Panyabungan. The Paired Sample t-Test result ($t(26) = 2.404$, $p = 0.024$) shows that the increase in students' interest scores following the implementation of Wordwall is statistically significant, and the effect size close to the medium boundary (Cohen's $d = 0.463$) indicates that this increase is also practically meaningful. Because this study used a one-group pretest-posttest design without a control group, this finding is interpreted as an increase in interest that occurred following, and is associated with, the use of Wordwall, rather than definitive proof that Wordwall was the sole cause of the increase. With that caveat, the finding supports the alternative hypothesis (H_a) proposed in Chapter II and is consistent with the theoretical prediction that Wordwall activities, through their novelty, interactivity, and immediate feedback, can trigger and maintain students' situational interest, as described in the Four-Phase Model of Interest Development (Hidi & Renninger, 2006; Renninger & Hidi, 2016).

The increase in the number of students in the High interest category, from 8 to 10 students, together with the decrease in standard deviation from 6.878 to 4.966, suggests that the treatment did not only raise the average level of interest but also reduced the gap between the most and least interested students. This pattern is broadly consistent with the theoretical expectation of the Four-Phase Model, in which repeated positive exposure to an engaging learning activity, delivered consistently across six instructional meetings, is theorized to help sustain students' interest over time (Hidi & Renninger, 2006). However, since this study only measured students' overall interest before and after the treatment rather than tracking it longitudinally across each of the six

meetings, this finding cannot directly confirm that individual students actually moved from a state of triggered situational interest to maintained situational interest; it only shows that interest, in aggregate, was higher and more evenly distributed after the treatment than before.

It is also worth noting how close the post-questionnaire mean came to crossing into the High category: at 60.26, it is only 0.74 points below the High boundary of 61 established in Table 3.7, meaning that a very small further shift, on the order of a fraction of a point per student, would have moved the overall sample mean into the High interest range. This proximity strengthens the practical reading of the effect size reported in Section E, but it should be interpreted cautiously given the design limitations discussed below. An examination of individual-level change scores also revealed that 8 of the 27 students (29.6%) recorded a lower total interest score in the post-questionnaire than in the pre-questionnaire, with decreases ranging from 1 to 12 points. Profiling these 8 students by their pre-questionnaire scores shows that all of them started in the High category or the upper end of the Moderate category, with a mean pre-questionnaire score of 62.75, compared to a mean of 54.47 among the 19 students whose scores stayed the same or increased.

This pattern is consistent with regression toward the mean, a statistical tendency for extreme initial scores to move closer to the group average on a second measurement independent of any treatment effect, and with a ceiling effect at the level of individual students: those who already reported high interest before the treatment had comparatively little room to score higher on a bounded Likert-scale instrument, and some may have reported more moderated or realistic self-assessments on the second administration simply because the novelty of completing the questionnaire itself, an instance of the Hawthorne effect in which participants' responses on a first measurement are inflated by the awareness of being observed or studied, had worn off (Adair, 1984; McCambridge et al., 2014). Because this study used a one-group pretest-

posttest design without a control group, it is not possible to statistically separate a genuine ceiling effect, a Hawthorne-related regression in self-report, and a true absence of benefit for already-interested students from one another using the present data; all three remain plausible, non-exclusive explanations for the decline observed in this subgroup, and this ambiguity is revisited as a limitation in Section F.5 and Chapter V.

2. Answering the Second Research Question

The second research question asked to what extent Wordwall affects each indicator of students' interest, namely attention, enthusiasm and participation, curiosity, and enjoyment. This question is answered both descriptively, using the pre-questionnaire and post-questionnaire mean scores of each indicator, and inferentially, using a separate Paired Sample t-Test per indicator, both processed using IBM SPSS Statistics. The descriptive comparison is presented first in Table 4.8, followed by the inferential test results in Table 4.9.

Table 4.8 Descriptive Comparison of Interest Indicators (Pre- and Post-Questionnaire)

Indicator	Pre Mean	Pre SD	Post Mean	Post SD	Mean Difference
Attention	13.52	1.909	14.74	1.347	+1.22
Enthusiasm and Participation	13.85	2.070	14.59	1.647	+0.74
Curiosity	14.22	1.968	14.44	1.577	+0.22
Enjoyment	15.33	2.572	16.48	2.408	+1.15

Source: IBM SPSS Statistics output, processed by the researcher (2026).

Table 4.8 shows that the four indicators of students' interest were not affected uniformly by the Wordwall treatment. Descriptively, Attention showed the largest increase in mean score (+1.22), followed

by Enjoyment (+1.15), Enthusiasm and Participation (+0.74), and Curiosity, which increased only slightly (+0.22). The mean of the four post-questionnaire indicator means ($14.74 + 14.59 + 14.44 + 16.48 = 60.25$) is consistent, within rounding, with the overall post-questionnaire mean of 60.26 reported in Table 4.3.

Table 4.9 Paired Samples Test per Indicator (Post-Questionnaire – Pre-Questionnaire)

Indicator	Mean Difference	SD	t	df	Sig. (2-tailed)	Cohen's d
Attention	1.22	1.450	4.380	26	0.000	0.843
Enthusiasm and Participation	0.74	2.330	1.652	26	0.111	0.318
Curiosity	0.22	2.207	0.523	26	0.605	0.101
Enjoyment	1.15	3.278	1.820	26	0.080	0.350

Source: IBM SPSS Statistics output, processed by the researcher (2026).

To answer the second research question more rigorously, a separate Paired Sample t-Test was additionally conducted for each of the four indicators, after the Shapiro-Wilk test confirmed that the difference scores for all four indicators were normally distributed (Attention: $W = 0.954$, $p = 0.272$; Enthusiasm and Participation: $W = 0.975$, $p = 0.738$; Curiosity: $W = 0.964$, $p = 0.457$; Enjoyment: $W = 0.974$, $p = 0.701$), satisfying the assumption required for the parametric test. As shown in Table 4.9, Attention was the only indicator for which the increase reached statistical significance, $t(26) = 4.380$, $p < 0.001$, with a large effect size (Cohen's $d = 0.843$). The increases in Enjoyment ($t(26) = 1.820$, $p = 0.080$, $d = 0.350$), Enthusiasm and Participation ($t(26) = 1.652$, $p = 0.111$, $d = 0.318$), and Curiosity ($t(26) = 0.523$, $p = 0.605$, $d = 0.101$) were positive in direction but did not reach statistical

significance at the 0.05 level, and their effect sizes ranged from small to negligible.

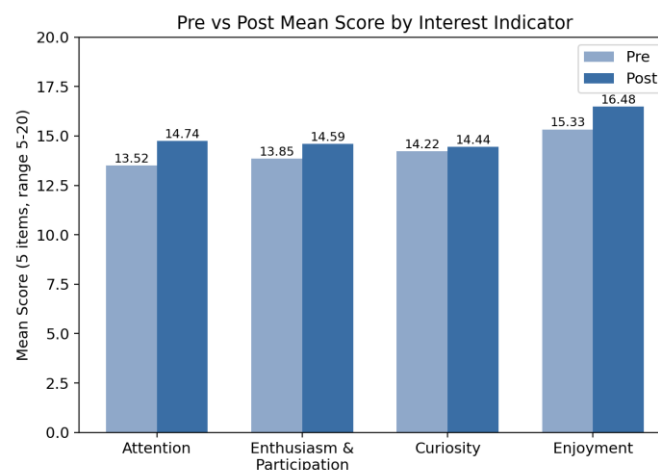
This pattern differs from what a purely descriptive reading of the mean differences alone would suggest, and it refines the interpretation of the second research question: rather than Wordwall producing a broadly distributed affective effect, its statistically reliable influence in this study is concentrated on Attention. This finding can be read through the Four-Phase Model of Interest Development (Hidi & Renninger, 2006; Renninger & Hidi, 2016). The earliest phase of the model, triggered situational interest, is fundamentally an orienting and attentional response, elicited when an environment presents novel, unexpected, or interactive stimuli. Wordwall's game mechanics, timed rounds, instant scoring, sound and visual feedback, and turn-based competition, are designed precisely to capture and hold this kind of moment-to-moment attention, which may explain why Attention produced the largest and only statistically significant change: the treatment's design maps most directly onto this specific phase of the model.

Enjoyment, by contrast, showed the second-largest descriptive increase (+1.15) but did not reach significance, plausibly because enjoyment is a more dispositional, affective response that varies more idiosyncratically across individual students ($SD = 2.408\text{--}2.572$, the highest of the four indicators in both measurements) than a shared attentional response to a structured game activity, so a consistent group-level effect is harder to detect with a sample of 27. Curiosity showed the smallest and least significant change, which is consistent with a ceiling effect at the indicator level: Curiosity already had a comparatively high pre-questionnaire mean (14.22 of a possible 20, or 71.1%), leaving the least room for measurable improvement, whereas Attention had the lowest pre-questionnaire mean among the four indicators (13.52, or 67.6%) and therefore the most room to rise. Because a separate

inferential test was not part of the originally planned analysis, this indicator-level testing is reported here as an additional analysis conducted to more fully answer the second research question, consistent with Deep et al. (2024), who similarly found that gamification's effects are not uniform across every dimension of engagement.

It should also be noted that four separate Paired Sample t-Tests, one per indicator, were conducted without applying a correction for multiple comparisons (such as a Bonferroni adjustment), which would ordinarily be used to control the family-wise Type I error rate when several hypotheses are tested on related outcomes from the same sample. This indicator-level testing was not part of the original single-hypothesis design of this study; it is reported here as an exploratory, hypothesis-generating analysis conducted to answer the second research question more fully, rather than as a confirmatory test. Consequently, the Attention finding ($p < 0.001$) is robust enough that it would likely remain significant even under a conservative correction (e.g., a Bonferroni-adjusted alpha of 0.0125 for four comparisons), but readers should interpret the exploratory indicator-level results, and any non-significant findings among them, with corresponding caution rather than as four independent confirmatory tests.

Figure 4.3 Pre- and Post-Questionnaire Mean Score by Interest Indicator



3. The Practical Magnitude of the Effect

The third research question asked how large the practical effect of Wordwall is on students' interest in learning English. As reported in Section E of this chapter, the effect size obtained (Cohen's $d = 0.463$) falls within the Small category, close to the boundary of the Medium category, according to Cohen's (1988) benchmark. This finding provides a more complete picture than the significance test alone, since statistical significance indicates only that the observed difference is unlikely to have occurred by chance, while the effect size indicates the magnitude of that difference in a way that is independent of sample size (Field, 2018).

4. Relation to Previous Studies

The findings of this study are broadly consistent with the three previous studies reviewed in Chapter II. Sakkir et al. (2023) found a significant improvement in EFL students' vocabulary mastery after the use of Wordwall, and reported that students also perceived Wordwall as making English lessons more interesting and enjoyable, a pattern that aligns with the significant increase in affective interest scores found in the present study, even though the earlier study measured vocabulary achievement rather than interest directly. Dindar et al. (2021) reported that gamified cooperative and competitive learning conditions produced greater engagement than conventional instruction, which is consistent with the present finding that game-based activities can generate a measurable, positive affective shift. Deep et al. (2024) found that gamification created a low-anxiety, stimulating learning environment that enhanced enjoyment and reduced psychological barriers to participation, a mechanism that is consistent with the increase in enjoyment-related items implied by the overall rise in interest scores observed in this study.

At the same time, the present study extends these previous studies by examining students' interest as a distinct affective construct,

measured through a validated Likert-scale instrument, within a vocational high school context in Indonesia, a combination that had not previously been investigated, as identified in the research gap discussed in Chapter II.

5. Limitations Affecting Interpretation

Several limitations should be considered when interpreting these findings. First, this study used a one-group pretest-posttest design without a control group, which means that internal validity threats such as history, maturation, and testing effects, discussed in Chapter I and Chapter III, cannot be fully ruled out. Second, the sample size of 27 students, while adequate for a classroom-based study, is relatively small and limits the extent to which the findings can be generalized beyond SMK Negeri 1 Panyabungan. Third, since both the pre-questionnaire and post-questionnaire instruments relied on students' self-reported responses to a Likert-scale questionnaire, the results may be subject to common method bias (Podsakoff et al., 2003), such as social desirability in reporting more positive attitudes after an enjoyable activity. These limitations are discussed further, together with corresponding suggestions, in Chapter V.

6. Practical Constraints During Implementation

Beyond the design-related limitations described above, two practical constraints were encountered during the implementation of the treatment and should be considered when interpreting the results. First, SMK Negeri 1 Panyabungan has only five projector (infocus) units shared across its 27 classrooms, and these units were also needed by other classes and subjects at the same hours, which meant that the researcher could only secure projector access for two of the six treatment meetings through advance scheduling with the school's facility coordinator. In the remaining four meetings, the Wordwall activities were carried out through students' personal smartphones without a shared classroom screen. This constraint was not something the

researcher could fully control, since throughout all six meetings the students' regular English teacher remained the instructor of record and continued to lead the core classroom instruction, so as not to disrupt students' usual learning routine; the researcher's own role in each session was limited to administering the Wordwall treatment segment under the teacher's supervision, rather than taking over the lesson to secure more consistent access to shared equipment.

This inconsistency in display medium across meetings may have introduced some variation in how collectively engaging each session felt, since a shared screen tends to support more visible, whole-class participation than individual smartphone access. Second, attendance varied slightly across the six meetings, as a small number of students were occasionally absent or unwell on a given meeting day, so not every one of the 27 students received exactly the same number of treatment exposures. Both constraints reflect common practical realities of conducting classroom-based research in a vocational high school setting with limited shared technology, and they are acknowledged as factors that may have introduced minor, uncontrolled variation into the treatment alongside the internal validity threats already discussed in Chapter III.

The projector inconsistency is also relevant to interpreting the Attention finding reported above: attentional engagement is typically supported by a shared visual focal point, so it might be expected that the two projector-based meetings produced a stronger attentional effect than the four smartphone-based meetings. Because this study measured interest only once before and once after the entire six-meeting treatment rather than after each individual session, it cannot determine whether the significant increase in Attention was driven mainly by the projector sessions, by the smartphone sessions, or fairly evenly by both; it is equally plausible that the personal, self-paced nature of smartphone-based Wordwall play sustained individual attention on its own,

independent of a shared screen. This is noted here as an open question for the session-by-session research design suggested in Chapter V, rather than a conclusion that can be drawn from the present data.

