CHAPTER III
RESEARCH METHOLODOGY

A. Research Design

The research methodology chosen for this study was a quasi-experimental design, designed to analyse the impact of game-based learning (GBL) integrated with the digital age (AI) on students' understanding of grammar. The research design employed a nonequivalent control group design, in which both groups underwent pre-tests and post-tests to measure improvements in grammar understanding after the implementation of GBL in the experimental group, in comparison to the conventional learning method in the control group. In this study, a comparison was made between the experimental and control groups; however, it should be noted that the selection and placement of these groups was not conducted via a randomisation process (Emzir, 2015)

The selection of these classes employed purposive sampling, wherein the researcher deliberately chose the two classes based on criteria of similarity in the number of students and relatively balanced student ability characteristics. The objective of this selection was to minimise initial differences between the two groups, thereby ensuring the validity of the research results. The purposive sampling technique is a nonrandom technique that does not need underlying theories or a set number of participants (Etikan, 2016).

This study which focuses specifically on SMAN 3 Bengkulu City, adds depth to the research by providing a detailed examination of the application of AI in real-world education, making the findings relevant to education stakeholders aiming to improve English language teaching in similar contexts. The combination of GBL and AI in grammar instruction aims to create an engaging learning environment that enhances students' interaction with the language.

The use of games in education, often referred to as serious games, has been widely discussed in educational literature due to their potential to enhance

motivation, participation, and information retention (Plass et al., 2015). By integrating AI tools, such as intelligent tutoring systems (ITS) or AI-based feedback mechanisms, students can receive personalised learning experiences tailored to their individual needs and performance, thereby enhancing the effectiveness of *PBL* GBL methods.

Table 3.1

Research De*sign*

Group	Pre Test	Treatment	Post Test
Experimental Class	O0	X	O1
Control Class	O0	O	O1

Description:

X : Use of media

O : Not using media X

O1 : Post test

O0 : Pre test

B. Research Variables

In this study, there are two types of variables: dependent variables, which depend on other variables or are influenced by other variables, and independent variables, which do not depend on other variables or affect other variables. This study's factors include:

1. Game Based Learning (Y)
2. Grammar Understanding (X)

C. Research Location and Time

1. Location of The Research

This research was conducted at SMAN 3 Bengkulu City, a public senior high school located in Bengkulu, Indonesia. The school was chosen not only

for its geographical accessibility but also for its diverse student body, representing a broad spectrum of academic abilities, socio-economic backgrounds, and language proficiencies. SMAN 3 has been identified as a school that is open to implementing digital and innovative teaching approaches, making it a suitable site for exploring the integration of AI in classroom based learning environments.

Bengkulu City itself, being part of a province with a rapidly growing education sector, provides a relevant context for this study. The city's commitment to improving educational outcomes aligns with the goals of this research, which seeks to explore how technological innovations can enhance learning processes. In the past, SMAN 3 has participated in various digital learning initiatives, thus positioning it as an ideal site for implementing GBL combined with AI.

2. Time of The Research

The study was place during the second semester of the 2024/2025 academic year, 11th grade students. Exactly on April 10th – Mei 10th 2025. This age group was selected because they are at a critical stage in their English language development, preparing for their national exams while also developing more advanced language skills. Additionally, they have sufficient prior knowledge of grammar, which allows for a more accurate measurement of the effects of the intervention.

The classroom environment at SMAN 3 is conducive to technology enhanced learning. Each classroom is equipped with computers and internet access, allowing for the seamless integration of digital tools such as AI based applications. The availability of these technological resources is a key factor in the success of this research, as the use of AI in GBL requires students to interact with digital platforms for real-time feedback and adaptive learning.

D. Population and Sample

The population of this study consists of all 11th grade students at SMAN 3 Bengkulu City of the 2024/2025 academic year. This population was selected because they have reached a level of English proficiency where they are familiar with basic grammar rules but may still face challenges in mastering more complex structures. The Number of Population are 457 students. According to the national curriculum, students at this level are expected to demonstrate competence in various aspects of grammar, including tenses, sentence structures, and subject-verb agreement, which are the primary focus areas of this study. The sampling method used in this study is purposive sampling, which involves selecting specific classes based on predetermined criteria. The two classes chosen for this study was similar baseline characteristics to control for extraneous variables. The criteria for selecting the classes include:

1. English proficiency level:

   Pre-test scores was used to ensure that both the experimental and control groups have a similar level of grammar proficiency at the start of the study.
2. Attendance and participation:

   Students with regular attendance and active participation in English classes was selected to ensure that they can complete the intervention without interruptions.
3. Willingness to participate:

   Students informed about the study and given the option to consent to participate. Only those who agree to participate was included in the final sample.

Table 3.2

Population of the study

Class	Male	Female	Total
A	15	27	42

B	14	26	40
C	20	21	41
D	16	24	40
E	17	23	40
F	13	26	39
G	14	18	32
H	19	13	32
I	20	15	35
J	19	33	42
K	20	14	34
L	18	22	40
Total			457

SMAN 3 Bengkulu City (2024/2025)

The final sample size wasconsist of approximately 63 students, with 32 students in the experimental group and 31 students in the control group. This sample size is sufficient for statistical analysis, allowing for meaningful comparisons between the two groups. According to Cohen (1988), a sample size of 32 per group is typically adequate for detecting medium to large effects in educational interventions (Brydges, 2019). This is very suitable with the research that was carried out in this study.

Table 3.3

Sample of the *study*

No	Group	Class	Male	Female	Total
1	The Experimental Class	XI H	19	13	32
2	The Control Class	XI G	13	18	31
Total					63

SMAN 3 Bengkulu City (2024/2025)

E. Research Instruments

A measuring instrument is a tool used to collect data and provide quantitative results. In this study, researchers used grammar-based multiple-choice tests as the main measuring instrument to measure students' understanding of English grammar, particularly in the use of present, past, and future tenses. The researcher administered pre-tests and post-tests to the experimental group and control group to determine the effectiveness of the Game-Based Learning (GBL) approach.

Grammar Test

Exams are an important part of the learning process. Exams consist of a series of structured questions designed to measure students' knowledge and understanding of a particular subject. In this study, the exam consisted of 50 multiple choice questions focusing on grammar topics in line with the English language curriculum, including simple present tense, simple past tense, and simple future tense.

A pre-test was administered before the intervention to measure students‘ initial grammar knowledge. After fours intervention sessions using Game Based Learning in the experimental group and conventional methods in the control group, a post-test was conducted to identify improvements in students’ grammar understanding.

The test questions were developed based on the grammar indicators listed in the curriculum and verified by English teachers and experts to ensure the reliability and validity of the content. Assessment was based on the number of correct answers, with each correct answer awarded one point. The total score was then calculated to determine each student's average score.

The difference in pre-test and post-test results between the two groups helped researchers evaluate the effectiveness of GBL as a grammar teaching strategy. Statistical analysis using an independent samples t-test was conducted to determine the significance of the results.

F. Data Collection Procedure

The data collection process was carried out in several stages to ensure a systematic and comprehensive approach to gathering the necessary information:

1. Preparation Stage:

   Before the data collection begins, several preparatory steps was taken:

   a) The pre-test and post-test was designed and validated by a panel of English language experts. The tests was piloted with a small group of students who are not part of the sample to ensure their reliability and validity.
   b) GBL materials have been prepared, and artificial intelligence (AI) tools to be used during the intervention have been selected. These tools may include grammar-based games, interactive quizzes, and intelligent tutoring systems that provide real-time feedback to students, such as *Kahoot!*.
   c) The school administration was contacted to gain formal permission to conduct the research. A detailed explanation of the study's purpose and procedures was provided to both the school authorities and the participants, ensuring transparency and voluntary participation.

2. Pre-test :

   Prior to the intervention, a pre-test was given to both the experimental and control groups to determine their initial level of grammar proficiency. This baseline measurement is crucial for making accurate comparisons with the post-test results. The pre-test wastake place in a controlled classroom environment, where students was required to complete the assessment within a specified time limit.

3. Implementation of Game-Based Learning:

After the pre-test, the experimental group wasgin a series of lessons incorporating game based learning strategies enhanced with *Kahoot!*. The GBL sessions wasinclude interactive grammar games designed to engage students in active learning over fours sessions. Learning is thus conceived as a four-stage session (Kolb, 1984). *Kahoot!* wasprovide immediate feedback on their answers, allowing them to correct mistakes and understand grammatical rules in real-time.

The control group, on the other hand, wasreceive *PBL* grammar instruction following the school's curriculum. This instruction wasprimarily involve teacher led explanations, textbook exercises, and drills without the use of any game based or AI enhanced activities.

4. Post-test :

Following the completion of the intervention, both the experimental and control groups wastake a post-test designed with a structure similar to that of the pre-test. The results of the post-test wasthen be compared to the pre-test scores to evaluate the extent of improvement in students' grammar proficiency.

G. Validity and Realibility

To ensure the validity and reliability of the research instruments, several steps was taken:

1. Validity:

The pre-test and post-test was designed in collaboration with English language experts to ensure they cover the key aspects of grammar that are relevant to the students' level. The tests was reviewed for content validity, ensuring that the questions accurately reflect the learning objectives outlined in the curriculum. Moreover, the questionnaires and

interview questions was reviewed to ensure they effectively capture students' attitudes and experiences with GBL and AI.

2. Construct Validity:

The construct validity of the tests was assessed by ensuring that they measure the theoretical constructs they are intended to assess in this case, students' grammar understanding. Items was carefully selected to represent different facets of grammar knowledge, such as understanding of rules, application in sentence structure, and error correction. To statistically test the validity of the instrument, the product moment correlation formula proposed by Pearson was used. This formula helps assess whether there is a significant relationship between the variable use of Game Based Learning on Digital era and the results of *PBL* Method. The formula used is as follows:

$$rxy = \frac{N \sum xy - (\sum x)(\sum y)}{\sqrt{(N \sum x^2 - \{\sum x^2\})\,(N \sum y^2 - \{\sum y^2\})}}$$

Notes:

N = Number of respondents

Rxy= product moment correlation coefficient

$\sum x$ = Number of item scores (x)

$\sum y$ = Number of variable scores (y)

$\sum xy$ = Number of multiplication of items (x) and variable scores (y)

$\sum x^2$ = Sum of squares of item scores (x)

$\sum y^2$ = Sum of variable scores (y)

3. Reliability

To ensure the reliability of the instruments, the pre-test and post-test was piloted with a small group of students before being administered to the full sample. The internal consistency of the test

items was assessed using Cronbach's Alpha, a statistic commonly used to measure the reliability of test scores. A Cronbach's Alpha value of 0.7 or higher wasindicate acceptable reliability. Similarly, the questionnaires was piloted to ensure that they are clear and understandable, and any ambiguous questions was revised before the final administration.

There are two types of reliability, namely internal reliability and external reliability. This research uses internal reliability test, because the calculation of reliability is done by analyzing the measurement data without involving external factors. The internal reliability test ensures that all items in the measurement instrument work consistently in measuring vocabulary mastery.

Various techniques for calculating reliability include using the Sperman Brown, Flanagan, Rulon, K-R 20, Hoyt, and Alpha Cronbach formulas. In this study, Alpha Cronbach formula was used to calculate the reliability of the measurement instrument. This formula was used because the instruments used in the study were in the form of scales or tests that measured vocabulary mastery on an ongoing basis.

The Alpha Cronbach formula used is as follows:

$$\alpha = \frac{k}{k-1}\left(1 - \frac{\sum \sigma_i^2}{\sigma_t^2}\right)$$

Description:

α = Cronbach's Alpha (reliability coefficient)

k = number of items (e.g., number of questions)

σ_i^2 = variance of each individual item

σ_t^2 = variance of the total test scores

$\sum\sigma_i^2$ = sum of the variances of all individual items

Table 3.4

Realibily Test

Case Processing Summary

		N	%
Cases	Valid	50	100.0
	Excluded[a]	0	.0
	Total	50	100.0

a. Listwise deletion based on all variables in the procedure.

Reliability Statistics

Cronbach's Alpha	N of Items
.736	60

H. Data Analysis

This researcher was employed quantitative methods for data analysis. The quantitative analysis focus on examining numerical data obtained from the grammar pre-tests and post-tests administered to both the experimental and control groups.

The primary objective is to evaluate the effectiveness of game based learning (GBL) enhanced by digital era (Artificial Intelligence) in improving students' grammar comprehension, in comparison to conventional teaching approaches.

1. Descriptive Statistics

   Descriptive statistics was used to summarize and describe the basic features of the data, providing an overall understanding of students' performance. Key statistical measures to be calculated include

a. Mean Scores:

The average score of both the experimental and control groups on the pre-test and post-test calculated. This provide insight into the overall grammar proficiency of the students in both groups before and after the intervention.

b. Standard Deviation (SD):

The standard deviation (SD) was utilized to assess the degree of variation or dispersion in scores within each group. It provides insight into how consistent the students' performance is. A lower standard deviation indicates that most students achieved similar scores, whereas a higher standard deviation reflects a wider range of performance levels among the students..

c. Minimum and Maximum Scores:

These values wasillustrate the performance range by identifying the minimum and maximum scores achieved by students in each group on both the pre-test and post-test.

d. Score Improvement:

The difference between pre-test and post-test scores was calculated for each student. The average improvement for each group was compared to assess how much students benefited from their respective instructional methods.

2. Normality Test

A normality test was performed to determine whether the distribution of grammar scores in both the experimental and control groups conformed to a normal distribution. This step is essential to validate the use of parametric statistical methods, such as the independent samples t-test, in the analysis.

The Shapiro-Wilk test was employed by the researcher, as it is considered more appropriate for small to medium sized samples. The statistical analysis was carried out using SPSS software, with a significance level set at 0.05. If the significance value (Sig.) exceeds

0.05, the data are considered to follow a normal distribution. However, if the value falls below 0.05, it indicates that the data are not normally distributed.

3. T-test

The SPSS t-test used in the analysis, with the proper formula applied. The experimental and control groups' post-test results was compared using an independent samples t-test. The purpose of this exam is to ascertain whether the two groups' reported differences in grammar proficiency are statistically significant.

Assuming that all other independent variables stay constant, the test is performed to assess the significance of each independent variable's partial effect on the dependent variable.

a. Null Hypothesis (H0) :
   There is no significant difference in post-test scores between the experimental and control groups.
b. Alternative Hypothesis (H1) :
   There is a significant difference in post-test scores between the experimental and control groups.

A p-value of less than 0.05 wasindicate a significant difference in the grammar understanding of students in the two groups, suggesting that GBL with AI was more effective than *PBL* methods.