CHAPTER III
METHODOLOGY

A. Approach and Type of The Research

In this study, the researcher used a quantitative method. Quantitative research was research conducted to determine phenomena by collecting data and then measuring it using statistical techniques. This definition was also supported by Creswell who said that quantitative research was an approach used to test objective theories by examining the relationships between variables or differences between groups. The data obtained from these variables were numerical in nature, allowing for analysis through statistical techniques (Creswell John, 2023). This method stemmed from the physical and chemical sciences that dealt with numbers in the form of scores and values, and this method was also called the data analysis method. It was a method suitable for collecting data, which was typically numerical, and this research employed mathematical models to analyze the data (Mohajan, 2020). Therefore, quantitative research was based on theory, presented using numbers, and analyzed using statistical techniques. The objective of quantitative research was to measure and analyze variables in order to obtain accurate results.

The research used an experimental research approach which was a complete data collection method because it could use other methods, such as tests, comparisons, and so on, to obtain data about the subject. This definition was also supported by Creswell who stated that the experimental method was a method that aimed to determine whether a certain treatment had an impact on certain results by giving a treatment to a certain group without another group, then determining how the two groups got a score and then knowing the results (Creswell John, 2023). There were other definitions, namely that this experimental approach was a complex research method, comparisons, measurements, and experiments were used to obtain empirical data processing with theoretical knowledge methods, so that they were formulated as scientific empirical facts (Maros et al., 2023).

B. Place and Time of The Research

This research was located at SMP Negeri 14 Seluma, where the school was located in Karang Anyar Village, Semidang Alas Maras Sub-District, Seluma Regency, Bengkulu 38875. The choice of research location was based on the researcher commitment to facilitate access to research data collection.

C. Research Design

This study employed a quasi-experimental approach, incorporating both an experimental group and a control group, but the determination of its members was not done randomly (Wafiqah et al., 2024). This design was chosen because in the context of school learning, random class assignment was often not possible due to administrative limitations and predetermined class structures.

According to Stratton, this design could be used to assess the effectiveness of interventions in delivering information or developing new concepts (Stratton, 2019). Therefore, this study applied pre-test and post-test nonequivalent control group design. This design was used to evaluate changes in students' vocabulary acquisition before and after treatment. In its implementation, a pre-test was given to both groups (control and experimental) to measure students' initial ability in vocabulary acquisition. After that, the experimental group received a treatment involving the use of the digital bamboozle game, Meanwhile, the control group did not receive any treatment. After the learning process was completed, a post-test was administered to both groups to measure changes in learning outcomes and compare the effects regarding the treatment in the experimental group. By

using this design, researcher were able to identify differences in students' vocabulary mastery between the experimental group and the control group, so that the effectiveness of using the digital bamboozle game in students' vocabulary mastery could be tested objectively.

Table 1. Pre-Test and Post-Test Nonequivalent Control Group Design

Experiment	Y1	X	Y2
Control	Y1		Y2

Notes:

Y1 : Pre-Test for control and experimental class

X : Treatment

Y2 : Post-Test for control and experimental class

D. Population and Sample

1. Population

The next stage in this study was to determine the population or sample of students involved in the study. Population was the whole of a group or object that had certain criteria to be the object of research. This term was also supported by Sugiyono who said that population was the whole element that become the generalization area (Sugiyono, 2019). This mean the scope or overall subject or unit that the researcher would measure. Meanwhile, according to Sukmawati, population was a general term consisting of subjects /

objects with certain qualities or characteristics that were recorded by researcher to be studied and then evaluated (Sukmawati et al., 2023). The population of this study were students of class VIII SMP Negeri 14 Seluma totaling 152 which were divided into 5 classes, among others:

Table 2. Total Number of class VIII SMP Negeri 14 Seluma

No	Class	Total
1	8A	31
2	8B	31
3	8C	30
4	8D	30
5	8E	30
Total Amount		152

(Source: SMP Negeri 14 Seluma: 2024)

2. Sample

The sample was the smallest part that come from a certain number or characteristics of the population (Sugiyono, 2019). This sample was made with the aim of collecting data that was useful for testing hypotheses and providing findings that could be applied to a larger population. The sample used in this study was purposive sampling. According to Sugiyono,

purposive sampling wass a sampling process that was carried out with certain considerations (Sugiyono, 2019). In this case, the consideration of researchers and teachers was the basis for choosing experimental and control classes, taking into account the condition of the class and the needs of students who still had a tendency to be passive in learning and others. After that, as for the classes that could be used as samples for this study, namely:

Table 3. Number of Experimental and Control Class Students

No	Class	Male	Female	Total
1	Experimental group	14	16	30
2	Control group	15	15	30
Total				60

(Source: SMP Negeri 14 Seluma: 2024)

E. Operational Definition

Operational was a study of the explanation of the variables to be studied in a study until they became operationally significant so that they could be used as a measuring instrument to measure research (Taqwin, 2022). In this case there was an explanation of the variables employed in the investigation, specifically:

1. Independent variable (Bamboozle Game)

   Baamboozle was a learning tool that could be played for both online and offline learning. It was an interactive game that was played in the classroom by forming teams. Through accessible and entertaining educational material, the platform helped increase student engagement in learning (Luke Edwards, 2024).

2. Dependent variable (Vocabulary Mastery)

   Vocabulary was a group of words used in language teaching, especially in teaching English (Wahyudi et al., 2021). One had to enrich their vocabulary in order to understand English lessons and be able to communicate in the language.

F. Data Collection Technique

Data collection was the process of conducting research in which researcher carried out activities that were directly related to the research objectives by using information, either in the form of data or other formats (Pasionista Vianitati, 2021). Another understanding was that data collection techniques were tools used to obtain data in the field, here were four data collection techniques (Sahir, 2022).

1. Test

To obtain data for this research required an instrument, in this instrument what was needed was a vocabulary test. This test related to language ability to be receptive or productive as a whole (Arisman, 2021). The assessment was conducted to evaluate an individuals comprehension of specific skills (Sahir, 2022). In this case the tests carried out are Pre-Test and Post-Test:

a. Pre-Test

In this study, the pre-test was a test used in the process of delivering material with the aim of knowing how well the material to be taught had been understood by students (Magdalena et al., 2021). The pre-test was the most important part of data collection techniques because it was a technique used to measure students' initial ability in vocabulary skills. In this case, the pre-test was administered prior to the students received the bamboozle game treatment. This was done so that the researcher could find out the initial ability of the students without any influence from the treatment. The pre-test consisted of 27 multiple-choice questions. The researcher gave clear instructions to students so that they understood

how to do it and gave enough time so that they did not rush in doing the questions.

After that, the analysis of the pre-test results was carried out to see the initial vocabulary ability to be used as a comparison after the students underwent treatment and provided a basic picture of the average vocabulary ability possessed by the class, so that this could be used by researcher to see the improvement after the treatment was carried out later.

b. Post Test

The post-test is administered following the conclusion of the treatment or the completion of the learning process for a specific subject, with the objective of understanding how well students understood the material in the treatment taught later (Magdalena et al., 2021).

This post-test was carried out after getting treatment phases which aimed to see or measure how much the students' vocabulary had increased after receiving the bamboozle game treatment and to see the comparison with the pre-test conducted previously. The time needed to do this post-test was after completing the treatment phases. Students were administered a post-test following

treatment, consisting of a multiple-choice format with a total of 27 items. After that, the stage of analyzing the scores from the pre-test and post-test was carried out to see the difference in scores obtained by students and to find out the significant effect between students who used the digital bamboozle game and those who did not. This provided information about the effectiveness of the bamboozle game toward student vocabulary.

c. Treatment

Treatment was the variable that was compared in the experiment (Taqwin, 2022). This meant that the treatment in this experimental class was the use of digital bamboozle game. The treatment was implemented to see the effect on students' vocabulary mastery. This activity was in the form of playing bamboozle as a method for mastering students' vocabulary in the experimental class. This bamboozle treatment activity was carried out with the aim of seeing the effect of the bamboozle game in mastering the vocabulary of eighth grade students and to compare the treatment results with the pre-test and post-test results to find out the extent of the game in students' vocabulary mastery. Before starting the game, the researcher

gave instructions to students about how the game would take place and gave a brief explanation of the new vocabulary. In this game the researcher role was to provide treatment and supervise students during the game, and ensure students could focus and understood the vocabulary contained in the game.

2. Vocabulary Test

A test was a tool that aimed to measure a person's increase in a person's understanding of certain material (Taqwin, 2022). In this case, the material in question was vocabulary. So a vocabulary test was a tool used to measure a person's understanding of vocabulary, of which there were two types of vocabulary tests commonly used to assess vocabulary understanding, namely size tests and level tests (Stoeckel, T., McLean, S., & Nation, 2021). Vocabulary test was used to measure the effect of digital bamboozle game toward student vocabulary. This test was given to the experimental and control classes to see the extent of students' achievement in mastering vocabulary. Tests were administered at the commencement and conclusion of the meeting to facilitate a

comparison of student learning outcomes between the experimental and control groups. In this study, researcher gave a multiple choice test to students. The number of questions in the test was 27 items with a duration of 50 minutes. Student test data were collected and analyzed by researcher to obtain research results. Before giving treatment to the experimental class, The researcher initially administered a pre-test of 27 questions to both the experimental and control courses. These questions came from items that have been validated through trials. As a first step, researcher conducted try outs in different classes and schools, with the aim of obtaining valid and reliable items. The number of questions that researcher used to try out was 60 items. After analyzing the results of the try out, researcher got 27 questions that met the criteria of validity and reliability.

Then, researcher gave treatment to the experimental class. In the final stage of the study, researcher administered a post-test to the experimental and control classes. The pre-test was used to determine whether the two groups had an equal level of ability on the dependent variable before treatment was given. After that, the

experimental class was given treatment through the implementation of digital bamboozle game, while the control class was not given treatment. Then, the researcher gave a post-test to both classes. The purpose of this step was to compare the final scores of the two classes and determine the effect of the treatment given. The Score Interval for English Teaching Vocabulary was:

Table 4. Score Interval For English Teaching Vocabulary

The Rage of Score	Scores Category	
20-18	Excellent To Very Good	sophisticated range-effective word idiom choice and usage etc.
17-14	Good To Average	adequate range occasional errors of word/idiom form, choice, usage but meaning not obscured.
13-10	Fair To Poor	limited range frequent errors of word/idiom form, choice, usage etc.
9-7	Very Poor	essentially translation-little knowledge of English vocabulary.

Source : (J. B. Heaton, 1976)

3. Documentation

Documentation was an activity carried out in research to obtain data in the form of photographs, videos, and others related to research when learning activities were carried out. Not only that, other documentation was conducted to guarantee that the activities conducted were precise and to add information to the research. Sahir explained that there were two kinds of documentation forms, namely first, making data groupings. For example in this study, researcher documented the types of information needed such as student data, grade data, attendance data, and others. Then second, making the variables to be collected. For example, researcher document what was needed for independent variables, for example collecting documents from the use of bamboozle games and dependent variables, for example documenting student learning outcomes represented by assessment scores. This could be a reference for researcher to collect information that must be documented. (Sahir, 2022).

G. Research Instruments

Research instruments were measuring instruments used to collect data from research objects. Instruments in quantitative research were very important because they were used to measure the variables to be studied by researcher (Taqwin, 2022). The instrument used as a measuring tool in this study was a test. Tests were instruments used to collect data or information in the form of a person's knowledge and skills (Taqwin, 2022). In this study, the test used is a vocabulary test conducted to test whether the use of bamboozle game has an effect toward students' vocabulary or not. Researcher used the same test questions for both classes. The test employed of multiple choice questions comprising of pre-test and post-test was given to 2 cycles, namely the experimental class and the control class. In addition, the vocabulary test was adjusted to the mastery of the target vocabulary according to the curriculum and learning materials. The vocabulary test consisted of 27 multiple choice questions, which were taken and adjusted to the books and assignments used at school.

H. Data Analysis Technique

Data collection techniques were methods or approaches used by researcher to collect data (Pasionista Vianitati, 2021). In this study, the researcher used analysis

of variation as a tool for analyzing experimental and control data. Analysis of variation known as ANOVA. The t-test and f-test were used to determine the effect of digital bamboozle Game toward student vocabulary. ANOVA analysis was carried out when the normality test and homogeneity test had met the requirements and showed appropriate results. To find out the results of the calculated data, the researcher used SPSS version 28 for windows. The following was an explanation of the data analysis technique that had to be met in the analysis of variance, namely:

1. Validity of the Test

   Validity was an instrument used to measure how well the measuring instrument was used to measure what should have been measured. (Raden Vina Iskandya Putri, 2023). Validity focused more on the accuracy or consistency of the instrument in measuring the desired concept or variable. There were three aspects of validation related to measurement, namely measuring instruments, measuring methods, and measurements. These three aspects determined whether the results of the measurement were valid or not. The validation test was carried out by comparing r-count with r-table (Taqwin, 2022). Overall, data test calculations

were conducted utilizing SPSS version 28 computational resources.

In addition, researcher also analyzed the test results using computer statistics such as SPSS version 28. The test results consisted of 60 questions containing four vocabulary indicators such as nouns, verbs, adjectives, and adverbs. This test was conducted to determine whether there is a significant effect of using digital bamboozle game toward students' vocabulary.

Table 5. Instrument Validity

Variabel	Item	r tabel	r count	Result
	S1	0,361	0,071	Unvalid
	S2	0,361	0,246	Unvllid
	S3	0,361	0,509	Valid
	S4	0,361	0,412	Valid
	S5	0,361	0,172	Unvalid
	S6	0,361	0,223	Unvalid
	S7	0,361	0,336	Unvalid
	S8	0,361	0,346	Unvalid
	S9	0,361	0,157	Unvalid
	S10	0,361	0,378	Valid
	S11	0,361	0,341	Unvalid
	S12	0,361	0,358	Unvalid
	S13	0,361	0,331	Unvalid

	S14	0,361	0,473	Valid
	S15	0,361	0,389	Valid
	S16	0,361	0,291	Unvalid
	S17	0,361	0,052	Unvalid
	S18	0,361	0,266	Unvalid
	S19	0,361	0,288	Unvalid
	S20	0,361	0,338	Unvalid
	S21	0,361	0,403	Valid
	S22	0,361	0,256	Unvalid
SOAL	S23	0,361	0,331	Unvalid
	S24	0,361	0,234	Unvalid
	S25	0,361	0,301	Unvalid
	S26	0,361	0,417	Valid
	S27	0,361	0,147	Unvalid
	S28	0,361	0,427	Valid
	S29	0,361	0,530	Valid
	S30	0,361	0,175	Unvalid
	S31	0,361	0,193	Unvalid
	S32	0,361	0,294	Unvalid
	S33	0,361	0,494	Valid
	S34	0,361	0,362	Valid
	S35	0,361	0,453	Valid
	S36	0,361	0,294	Unvalid
	S37	0,361	0,460	Valid

	S38	0,361	0,266	Unvalid
	S39	0,361	0,467	Valid
	S40	0,361	0,120	Unvalid
	S41	0,361	0,123	Unvalid
	S42	0,361	0,388	Valid
	S43	0,361	0,149	Unvalid
	S44	0,361	0,537	Valid
	S45	0,361	0,080	Unvalid
	S46	0,361	0,479	Valid
	S47	0,361	0,376	Valid
	S48	0,361	0,.428	Valid
	S49	0,361	0,427	Valid
	S50	0,361	0,594	Valid
	S51	0,361	0,123	Unvalid
	S52	0,361	0,361	Unvalid
	S53	0,361	0,362	Valid
	S54	0,361	0,486	Valid
	S55	0,361	0,135	Unvalid
	S56	0,361	0,503	Valid
	S57	0,361	0,218	Unvalid
	S58	0,361	0,423	Valid
	S59	0,361	0,444	Valid
	S60	0,361	0,428	Valid

According to the results of the validity testing of the questions in the aforementioned

table, it is evident that out of the 60 statement items, the number of valid questions was 27 question items, because 33 items of the rcount statement were < r-table so that the statement was invalid.

Table 6. Test Instrument Validity Criteria

Cronbach's Alpa	Validity Level
0,81 – 1,00	Highly
0,61 – 0,80	High
0,41– 0,60	Fairly
0,21 – 0,40	Low
0,00 – 0,20	Very Low

Source: (Widodo et al., 2023)

In this validity test, researcher used multiple choice vocabulary questions totaling 60 questions. Then the researcher conducted a trial in one of the junior high schools that had the same class and accreditation as the school that the researcher would use for the place where the researcher conducted the research later. The trial was conducted with a total of 30 students as the test subject. After conducting the trial, the researcher processed the data on the trial questions by calculating the validity of each question. Therefore, researchers found a total of 27 valid

questions out of 60 questions, with the following indicators:

Table 7. Indicators of Vocabulary Test Validity

No	Indicators	Number of Item	Items
1	Understanding word meaning and pronouns in context	6	3, 10, 21, 26, 28, 29,
2	Identifying adjectives in descriptions	4	4, 14, 47, 50
3	Using verbs correctly	7	15, 33, 34, 35, 37, 39, 42
4	Understanding synonyms and antonyms	4	46, 49, 53, 54,56
5	Identifying nouns professions and roles	2	58, 59
6	Recognizing collocations in English	3	44, 48, 60

	Total	**27**

2. Reliability of the test

Reliability referred to the capacity of the measuring equipment to yield consistent data when it was administered at different periods (Taqwin, 2022). If a test or instrument was said to be reliable, the results were consistent every time it was used. The reliability value was assessed via Cronbach Alpha, if the Cronbach Alpha r value is ≥ 0.70 (70%) then the level of reliability is considered satisfactory (Pasionista Vianitati, 2021).

The meaning of the category of Cronbach's Alpha value was as follows:

Table 8. Test Instrument Reliability Categorization

Test Reliability Coefficient	**Categorization**
1,00 > 0,90	Very high reliability
0,75 - 0,90	Good reliability
0,50 - 0,75	Medium reliability
-1,00 < 0,50	Poor reliability

Source : (Hari Sugiharto Setyaedhi, 2024)

In this reliability test, researcher used multiple choice questions with a total of 60 questions. Then the researchers conducted a trial in one of the junior high schools that had the same level as the school that the researchers would use for the place where the researchers conducted the research later. The trial was carried out with the number of students as test subjects as many as 30 people. After conducting the trial, the researchers processed the test data by calculating the reliability. After processing the reliability data of the test questions, the reliability test coefficient was 0.853 which was based on the category above that the item was said to be reliable if it reached a value of 0.600 or higher (Hari Sugiharto Setyaedhi, 2024). Therefore, it can be assumed that this test had a good level of reliability.

Table 9. The Result of Reliability

Reliability Statistics	
Cronbach's Alpha	N of Items
0.853	60
Reliabel	

Based on the table above, it was concluded that the 60 questions tested on 30 students were very suitable to be used in the pre-test and post-test which would be carried out by researcher later

because they showed the results of Cronbach's Alpha reaching 0.853. which could be seen from table 7 with the Good reliability category.

3. Item Difficulty Test

Item difficulty test was a very important part of research instrument development, including to see the effect of digital bamboozle game toward eighth grade students' vocabulary acquisition. In this case, it refered to the level of difficulty of each question in the evaluation instrument. The test was conducted to ascertain the extent to which a question item could be answered correctly by the target population. In the context of the digital bamboozle game, item difficulty analysis was conducted to ensure that each item had an appropriate level of difficulty for students' abilities. This testing process involved collecting data from students who were given a series of kosataka questions in the digital bamboozle game. After students completed the exercise, the results obtained were further analyzed. The difficulty of an item was ascertained by dividing the number of students who answered it correctly by the total number of test participants.

Then there are testing criteria in the item difficulty test:

The general formula is:

$P = \frac{Rh + Ri}{Nh + Ni} \times 100\%$

Where:

P = difficulty level in percent

N_h = Number of test takers in high score group

R_h = Number of correct answers in high score group

N_i= Number of low score group test takers

R_i = Number of correct answers in low score group

$P = \frac{445}{8} + \frac{104}{8} \times 100\ \%$

P = 68,62%

Table 10. Interpretation item difficulty test

Criteria Score	**Interpretation**
0,21 – 0,40	Difficult Category
0,41 – 0,70	Medium Category
0,71 – 1,00	Easy Category

Source: (Ida Ayu, 2019)

Consequently, it was imperative for the researcher to assess the difficulty level of the items to ensure balanced complexity in the vocabulary questions used for the study. This was so that the questions were not too easy and therefore less

challenging, or too difficult and therefore caused students to feel pressured. An item difficulty analysis helped in ensuring that the materials used were appropriate to the needs and abilities of eighth grade junior high school students. So, based on the calculation of the level of difficulty in the questions given, it reached 68.62%, so based on the category of moderate difficulty 41%-70% it was classified as a moderate level of difficulty so that learning objectives and student abilities were the main target.

4. Item Discrimination Test

The Distinguishing Power Test was an important part of question instruments, including in this study which focused on the effectiveness of the digital bamboozle game toward students' vocabulary mastery. This test aimed to determine degree to which the itemd in the vocabulary test were able to differentiate between students who had high and low abilities in mastering English vocabulary. Distinguishing power referred to the ability of an item to identify differences between students who had a better understanding of vocabulary and those who were still lacking. Items with high discriminating power showed a clear

difference between students who had mastered more vocabulary and students who still had limitations in this aspect. The Differentiability Test process in this study was conducted by categorizing students into two groups based on their vocabulary test scores, upper group i.e. students with the highest score in the vocabulary test and lower group i.e. students with the lowest score in the vocabulary test. Generally, 27% of the highest-scoring students and 27% of the lowest-scoring students were sampled for the power difference analysis. The distinguishing power was computed by subtracted ratio of accurate responses from the lower group to the ratio accurate responses from the upper group.

The general formula is:

DP = $\frac{U-L}{N}$

Where:

DP = Question Distinguishing Power

U = Number of correct answers from High score group test takers

L = Number of correct answers from Low score group test takers

N = Number of students per group (27%)

DP = $\frac{445-104}{16}$

DP = 21,31%

Then there were testing criteria in the item discrimination test:

Table 11. Interpretation Item Discrimination Test

Item discrimination test	**Interpretation of result**
DP ≥0,40	Good
+0,21 < DP < 0,40	Fairly Qualified
0, < DP ≤ 0,20	Weak

Source: (Restiyawati, 2023)

Based on the results of the distinguishing power calculation in this study reached 21,31, when compared with the interpretation above, it is quite eligible, namely if + 0.21 < DP (21,31) < 0.40, it could be interpreted that the Distinguishing Power of this study was quite eligible.

5. Normality Test

The normalization test was a statistical assessment that evaluated if the sample data followed a normal distribution. The objective of this test was to ascertain the normality of the data distribution. This study employed a normality test to assess the data's normalcy. The formula used in this study as a data analysis tool indicated that the

data were normally distributed if the normal probability plot data results showed that the residual data formed a straight line or were close to straight, while the histogram results were indicated by a bell-shaped residual image (Ira Yuniati, 2022). Then there were testing criteria in the Normality Test, namely if the sig value was < 0.05 concluded that the data were not normally distributed, and if the value was > 0.05 concluded that the data were normally distributed.

6. Homogeneity Test

The homogeneity test was employed to ascertain whether the variances of multiple populations were the same. This test was typically conducted as an initial step prior to performing analyses such as the Independent Sample T-Test or ANOVA. One of the basic assumptions of ANOVA was if two groups were normally distributed (Ira Yuniati, 2022). To test whether the data were homogeneously distributed, a comparison was made between the variances of two or more groups of data. If the variances of the groups turned out to be the same, then the homogeneity test was no longer needed because the data were considered homogeneous. However,

this test could only be performed if the data followed a normal distribution. The purpose of the homogeneity test was to ensure that differences that appeared in parametric statistical tests (such as t-test, ANOVA, or ANCOVA) were really caused by differences between groups, not because of inequality within the groups themselves. Therefore, it was very important to conduct a homogeneity test before comparing two or more groups so that the results of the analysis truly reflected differences between groups, not on the basis of data that were not comparable.

7. T-test

T-test was a test used to compare the difference between the data analysis of the experimental data group and the control group (Sahir, 2022). The purpose of this T-test was to determine the significance between each data variable between students' vocabulary comprehension ability by using digital bamboozle game in the experimental class and students' vocabulary comprehension ability without using digital bamboozle game. The analysis was conducted using the T test on SPSS version 28, accompanied by the use of relevant formulas.

Assuming other independent variables remained constant, this test aimed to evaluate the significance of the partial effect of each independent variable on the dependent variable. With decision-making criteria, if the sig value (2-tailed) > 0.05, then Ho was accepted and Ha was rejected, and if the sig value (2-tailed) < 0.05, then Ho was rejected and Ha was accepted.

8. F- Test

   The F-Test was used to determine whether there was a significant difference or effect between the independent and dependent variables simultaneously (Sahir, 2022). The F test was conducted by comparing the significance value (sig) with the confidence level (α) of 0.05 or by comparing the calculated F value with the F table. The decision criterion for this F test was that if the significance value was greater than 0.05, then the independent variables did not affect the dependent variable. If the significance value was less than < 0.05, then the independent variables collectively influenced the dependent variable (Putri & Andriansyah, 2022). A comparison was also made between the calculated F value and the F table. If the calculated F value was greater than the F table

value, then Ho was rejected and Ha was accepted. The F table value was determined based on the degrees of freedom (df1; df2), or (k; n-k-1), where (k) was the number of independent variables and (n) was the number of samples used.

9. The Effect Size

As a final step in this analysis test to determine the extent to which the effective technique or media was used, then the researcher applied Cohen's formula 33. To calculate how much the effect size of the effect of the digital bamboozle game was calculated using the formula, as stated by Cohen was cited by Lis Darvia as follows (Darvia et al., 2024):

$$d=\frac{(mean\ of\ group\ A\ mean\ of\ Grup\ B)}{Pooled\ Standart\ Deviation}$$

Pooled Standart Deviation =

(standart deviation of group 1+ standart deviation of group 2)

Where the criteria of the effect size level are:

Table 12. Effect Size

0,00 – 0,195	Very weak effect
0,20 – 0,395	Weak effect
0,40 – 0,595	Modest effect
0,60 – 0,749	Strong effect

0,80 – 1, 00	Very strong effect

10. Uji Hypothesis

Then the last step was to test the hypothesis, which was to find out whether the alternative hypothesis (Ha) in the study was acceptable or not. In this research, there were Null Hypothesis (Ho) which stated there was no a significant effect of using Digital Bamboozle Game Toward Students' Vocabulary Mastery, and Alternative Hypothesis (Ha) which stated There was a significant effect of using Digital Bamboozle Game Toward Students' Vocabulary Mastery. Then there were testing criteria in the Hypothesis Test, namely:

- (Ha) There was a significant of use Digital Bamboozle Game Toward Students' Vocabulary Mastery if p-value was < sig α = 0,05. Then hyphothesis (Ha) accepted and the hyphothesis (Ho) rejected.
- (Ho) There was no a significant effect of use Digital Bamboozle Game Toward Students' Vocabulary Mastery if p-value was > sig α = 0,05. Then hyphothesis (Ha) rejected and hyphothesis (Ho) accepted.