

CHAPTER III

RESEARCH METHOD

A. Research Design

This study employed a quasi-experimental design without an equivalent control group. According to Sugiyono (2013), this design involves an experimental group and a control group selected without randomizing the subjects. Both groups take a pretest and a post-test, but only the experimental group receives the treatment. Thus, a quasi-experimental design was used to test the effects of an intervention even though the researcher cannot randomly assign subjects. This design is widely applied in educational research because researchers generally use existing classes, making it more feasible to conduct in a real learning environment (Septiana *et al.*, 2025).

Based on this concept, this study involved two groups namely an experimental group and a control group. In the initial stage, both groups took a pre-test to measure students' baseline vocabulary proficiency. Subsequently, the experimental group received an intervention involving the use of the Assemblr Edu app in learning, while the control group participated in learning using conventional methods without the app. After the intervention was fully implemented, both groups took a post-test to assess changes in students' learning outcomes.

Through this procedure, researchers can compare improvements in vocabulary mastery between the two groups while controlling for differences in initial ability using pre-test scores (Ekayana *et al.*, 2025). Furthermore, the Non-Equivalent Control Group Design is considered effective for testing the effect of the treatment in real-world learning conditions because it can still yield reliable findings through an analysis of the difference between pre-test and post-test scores, even without subject randomization (Okeke *et al.*, 2025). Therefore, this design is appropriate for testing the effect of using the Assemblr Edu app on students' vocabulary mastery. This research design can be seen in the following table:

Table III.4.1
Research design

Experiment Group	O1	X	O2
Control Group	O3		O4

(Sugiyono, 2013)

Description:

X: Assemblr Edu Application

O1: Pretest in experimental class

O2: Posttest on the experimental class

O3: Pretest in control class

O4: Posttest on control class

B. Location and Time of Research

This study was conducted from April 30 to May 30, 2026, at state Junior High School 20 Bengkulu, Bengkulu Province, during the second semester of the 2025–2026 school year.

C. Population and Sample

1. Population

The population in this study was all eighth-grade students at state Junior High School 20 Bengkulu City in the current academic year. This population was selected because eighth-grade students had already acquired the basics of English in seventh grade and were therefore considered ready to accept new learning technologies, including Augmented Reality (AR), as an aid in understanding vocabulary.

Table III.2
Population of the Research

No	Class	Male student	Female student	Total	Mean Score
1	A	20	12	32	77
2	B	18	12	30	83
3	C	16	14	29	77
4	D	17	11	28	79
5	E	16	14	30	79
6	F	17	11	28	80
7	G	18	11	29	84
Total				202	

2. Sample

This study used purposive sampling. Purposive sampling is a technique for determining samples based on specific considerations in accordance with the research objectives. According to Sugiyono (2013), purposive sampling is a sampling technique based on specific characteristics relevant to the research. In addition, purposive sampling is also defined as a sample selection technique by selecting individuals or groups who are considered most informative and relevant to the phenomenon being studied (Dewi et al., 2025). Frankel *et al.* (2012) also explain that in purposive sampling, researchers use personal judgment based on prior

knowledge of the population to select a sample deemed capable of providing the data required for the study.

In this study, personal judgment in sample selection was conducted by considering the students' average report card scores and recommendations from the English teacher. The students' average report card scores were considered to identify classes with relatively lower and comparable academic achievement. Based on the report card scores, Class A and Class G had the same average score of 77.

Therefore, these two classes were considered suitable and comparable for selection as the experimental and control groups. The English teacher's recommendation was also considered to ensure that the selected classes were appropriate for the research and had similar classroom conditions. The criteria for sample selection were as follows:

- a. The sample consisted of eighth-grade students at State Junior High School 20 Bengkulu City who were enrolled in the 2025–2026 academic year.
- b. The average report-card scores of the eighth-grade classes were used as a basis for sample selection, with the class having the lowest average score selected as the experimental class and the class having the highest average score selected as the control class.

- c. The final selection of the classes considered the English teacher's recommendations and the availability and permission of the school to conduct the research.

Top of Form

Bottom of Form

The selected samples of this study are presented in the following table.

Table III. 3
Sample of the Research

No	Class	Male student	Female student	Total
1	A	17	11	28
2	G	17	11	28
Total Sample				56

Initially, there were 32 students in Class A and 29 students in Class G. However, not all students could be included in the research sample because this study used pre-test and post-test data. In Class A, one student did not take the pre-test due to illness, while during the posttest, three students were absent: one with a valid excuse and two due to illness. Meanwhile, in Class G, there was one student who did not take the posttest with a valid excuse. Therefore, students who did not take one of the tests were excluded from the data analysis to ensure that the analysed data was

complete and the research results remained accurate. Thus, the final sample size used in this study was 56 students, consisting of 28 students from Class A and 28 students from Class G.

D. Operational definition of Variable

A research variable is an attribute, characteristic, or value of a person, object, or activity that has certain variations and is determined by the researcher to be studied and concluded (Sugiyono, 2013). Based on their function and role in research, variables are divided into independent variables and dependent variables. Independent variables are variables that influence or cause change, while dependent variables are variables that are influenced or result from independent variables. In this study, there are two main variables, namely the use of the Assemblr Edu application as an independent variable and vocabulary mastery as a dependent variable.

1. Vocabulary

Vocabulary is a collection of words known and used by a person in a language. Vocabulary plays an important role in language learning because it is the basis for developing language skills, such as reading, writing, speaking, and listening. The more extensive a student's vocabulary, the better their ability to understand and convey meaning in English communication.

2. Assemblr Edu

Assemblr Edu is an Augmented Reality (AR)-based learning application that provides interactive learning media in the form of three-dimensional objects. This application allows teachers and students to interact with digital visual content, making learning more interesting, contextual, and meaningful.

E. Data Collecting Technique

1. Instrument Pilot Test

Before the study was conducted, the researcher first conducted a pilot test of the instrument to assess the quality of the research instrument to be used. The pilot test aimed to evaluate the validity and reliability of the questions to ensure that the instrument was suitable for the study. The pilot test was conducted on May 6, 2026, at the same school as the research site but did not use the main research sample. The main research sample consisted of students from Class A and Class G, while the pilot test was administered to 30 students from Class B. This was done to ensure that the students in the main research sample were not influenced by questions that had already been tested. During the pilot test, the researcher administered 60 vocabulary questions.

2. Test

In the world of education and science, tests are often used as a tool to measure the extent to which a person understands the concepts, facts, or skills that have been taught. Tests serve as a means of evaluation to determine the level of learning achievement of students after participating in a particular learning process (Afrilianti *et al.*, 2025).

Tests can also be understood as a method of data collection conducted through a series of questions or problems that must be answered, responded to, or through tasks that must be carried out by the subjects being studied. Through tests, researchers can obtain quantitative data that describes the ability or mastery of students of the material that has been given (Anantasia, 2025).

In this study, tests were used as the main instrument to measure students' vocabulary mastery as a dependent variable. The tests were given to students in the form of pre-tests and post-tests to determine changes in students' vocabulary abilities before and after being given treatment using the Assemblr Edu application. The use of pre-tests and post-tests aims to measure the dependent variable before and after treatment is given, thereby the

effectiveness of the treatment applied can be determined (Pellas, 2023).

a. Pre-test

A pre-test is a test given to students before the treatment is carried out. The pre-test was conducted on May 11, 2026, to determine the students' initial ability in mastering English vocabulary, especially vocabulary related to the Environment. The pre-test results are used as a basis for determining the initial condition of the students and to ensure that the initial abilities of the experimental group and the control group are relatively comparable before the treatment is given.

b. Post-test

A post-test was a test administered to students after the treatment had been completed. The post-test was conducted on May 26, 2026, to determine the students' level of vocabulary mastery after the learning process using the Assemblr Edu application in the experimental group. The post-test results are then compared with the pre-test results to determine whether there has been an increase in the students' vocabulary skills. The comparison between the pre-test and post-test results is used to determine whether the use of the Assemblr Edu application has an effect on the students' vocabulary mastery.

c. Vocabulary Test

The vocabulary test in this study was designed in a multiple-choice format. The multiple-choice format was chosen to obtain objective, consistent data that could be easily analysed quantitatively. Each question had one correct answer and several distractors, thus students' vocabulary proficiency could be measured more accurately. The test measured students' ability to recognize and understand English vocabulary related to the topics taught during the treatment. The test items were developed based on three units from the teaching material: **Unit 1: Look Around You**, which covered *Saving the Environment* and *Saving Water*; **Unit 2: This Is the Way**, which focused on *Saving the Earth*; and **Unit 3: Act Now**, which covered *Washing Hands*. Through this test, students were expected to choose the correct meaning of a word and recognize the use of vocabulary in sentence contexts. The students' test scores were then classified based on their level of vocabulary mastery. The classification of students' scores is presented in table:

3. Treatment

a. The treatment of experimental class

The treatment was conducted over 4 meetings using the Assembler Edu application on the “Love Our World” material. Each meeting lasted 2 class periods (90 minutes) and utilized the discovery learning model. In the first meeting on May 12, 2026, from 10.25-11.45 a.m., students learned vocabulary related to clean and polluted environments. Students scanned AR markers to view 3D objects such as trees, clean air, trash and pollution. They then observe and categorized the vocabulary into “healthy environment” and “bad environment”. The activity continued with group discussions and practice expressing simple opinions about the environment. In the second meeting on May 18, 2026, from 13.00-14.50 p.m. students learned vocabulary related to the water usage and water saving habits. Through Assemblr Edu, students observed actions such as turning on a faucet, turning off a faucet. Students then completed worksheets and discussed the actions that qualify as water saving versus water wasting. In the third meeting on May 19, 2026, from 10.25-11.45 a.m., students learn to identify environment problems and simple solutions. Assemblr Edu display

conditions such as flooding, deforestation and solutions for protecting environment. Students discuss in groups to match problems with solution then present the results. The last meeting on May 25, 2026, from 13.00-14.50 p.m., students learn steps for washing hands and the use of imperative sentences.

b. The treatments of control class

The treatment in this study was implemented over 4 meeting using printed pictures on the “Love Our World” material. In the first meeting on 12 May 2026 from 07.30-08.50 AM, the researcher showed pictures of clean and dirty environments to introduce vocabulary such as clean, dirty, trash, pollution, and tree. Students observed the pictures, imitated the pronunciation of the vocabulary, and wrote simple sentences based on the pictures. In the second meeting on 18 May 2026 10.25-11.45 a.m., the researcher used images of water-saving activities to introduce vocabulary such as water, tap, save water, and waste water. Students repeated the vocabulary and completed simple sentences about water conservation. In the third meeting on May 19, 2026, from 07.30-08.50 a.m., the researcher showed pictures of forests, floods, and environmental damage to explain vocabulary such as forest, flood, and deforestation. The students then grouped the vocabulary into the categories

of good, problem, and solution. In the last meeting on May 25, 2026, from 10.25-11.45 a.m., the researcher used images of the steps for washing hands to introduce vocabulary such as wash, soap, rinse, and clean. Students sequenced the images of the handwashing steps and wrote simple instructions based on the provided images.

4. Documentation

Documentation is a data collection technique that can take the form of writing, images, or monumental works by an individual (Mukhtadibillah et al., 2020). Through documentation, researchers can obtain written and visual supporting data relevant to their research. In the context of educational research, documentation is used to collect data such as student lists, lesson plans, and photos of learning activities during the research process. This documentation data serves to complement and strengthen the main data obtained through tests.

F. Research instruments

Research instruments are used to collect data from research subjects. In quantitative research, instruments play a very important role because they are used to measure the variables being studied. The test in this study consists of 28

multiple choices develop based on indicators organized according to the table below.

Table III.5
Blueprint Test

Unit	Indicator	Word Class	Question number	Total
Topic: Look Around You - Saving the Environment & Saving water	Identifying vocabulary related to the environment and water	Noun	3, 7, 15	10
	Conservation.			
	Understanding verbs related to environmental conservation and water conservation	Verb	2, 14, 23, 28	
	Understanding adjectives related to environmental conditions and water conservation.	Adjective	13, 16, 19	

Topic: This is the way – Saving the Earth	Identifying vocabulary related to the global environment (Earth)	Noun	18, 20, 24	10
	Understanding verbs in caring for the Earth.	Verb	4, 5, 8, 21, 26	
	Understanding the meaning of messages in posters and simple texts about saving the earth	Adjective	17, 22	
	Identifying vocabulary related to handwashing	Noun	9, 10	
Topic: Act Now – Washing hand	Understanding verbs related to handwashing	Verb	12, 25	8
	Understanding adjectives	Adjective	1, 6, 11, 27	

related to
hygiene and
health

G. Data analysis techniques

1. Validity

Validity aims to determine whether the test instrument used is appropriate and suitable for measuring the variables under study. An instrument is considered valid if it is able to measure what it is supposed to measure in accordance with the research objectives. The researcher used expert validity and Validity tests-based SPSS.

a. Expert Validity

The validation process was conducted three times through consultation with expert in the fields of English language education. The first validation was conducted on April 27, 2026. At this stage, the expert provides several inputs and suggestions for improvements to the instrument that have been developed. The suggested improvements including adding clearer directions, adjusting answer A, B, C, D to better align with the theme and indicator of the questions, and increasing the number of questions from the initial 45 to 60 items. The second validation was conducted on April 30, 2026. After reviewing the results of the previous revisions, the experts

provided additional suggestions including the addition of an instrument header, a participant identification column covering name and class and time allocation guidelines. The third validation was conducted on May 4, 2026. At this stage, the revised instrument was re-examined by expert. Based on the assessment results, all previously provides suggestions and improvement had been properly implemented

b. Validity test-based SPSS

Validity testing was conducted using the Product Moment correlation by comparing the calculated r count with the r table at a 5% significance level with 30 students as respondents, resulting in a r table value of 0.361. An item was deemed valid if the calculated r table was greater than the r table value (0.361). Based on the calculations, 28 items were deemed valid because they had a calculated r count greater than 0.361, while the other 32 items were deemed invalid. Therefore, only the 28 valid items were used as the research instrument in Class A and Class G at Junior High School 20, Bengkulu. The result f the validity test for the research instruments is presented in the table below.

Table III.6
Validity Test Result

Validity test result	Question number	Total
Valid	3, 4, 5, 8, 10, 11, 12, 13, 22, 23, 27, 28, 30, 31, 33, 34, 37, 43, 44, 45, 46, 47, 48, 51, 52, 54, 55, 59.	28
Invalid	1, 2, 6, 7, 9, 14, 15, 16, 17, 18, 19, 20, 21, 24, 25, 26, 29, 32, 35, 36, 38, 39, 40, 41, 42, 49, 50, 53, 56, 57, 58, 60.	32
Total		60

The table shows that out of 60 test items, 28 items were declared valid and 32 items were declared invalid. Therefore, only the 28 valid items were used as the research instrument for the pre-test and post-test. The complete results of the validity test, including the item analysis and SPSS output, can be found in Appendix 11 on page 171.

2. The Level of difficulty test

Item difficulty testing is conducted to determine the extent to which each item in a vocabulary test has a level of difficulty that is appropriate for the students' abilities. Item difficulty indicates the proportion of students who can answer a question correctly. Good questions are neither too easy nor too difficult, therefore that they can accurately measure students' abilities and distinguish the level of vocabulary mastery among students. The difficulty levels are as follow:

Table III.9
Test Item Difficulty Level Category

Level of Difficulty	Description
$P < 0,20$	Hard
$0,20 \leq P \leq 0,80$	Currently
$P > 0,80$	Easy

Source (Riswana & Nasution, 2024)

Based on data from 30 student responses, the results of the item difficult level test showed that 55 question were rated as “currently,” while the remaining 5 were rated as “hard.”

3. Item discrimination test

The discrimination power of a question shows the ability of an item to distinguish between students with high

abilities and students with low abilities. Questions with good discrimination power will be answered correctly more often by students with high abilities than by students with low abilities. The interpretation of the discrimination power index was based on the established criteria presented below.

Table III.10
Interpretation of the power Difference test

T-test	Interpretation of results
$0,71 \leq DP \leq 1,00$	Very good
$0,41 \leq DP \leq 0,70$	Good
If $+0,21 < DP < 0,40$	Sufficiently qualified
If $0 < DP \leq 0,20$	Weak

Source: (Ningsih & Widowati, 2021)

Based on the result of the discriminant test, there was 1 item categorized as “very good,” 18 items as “good,” 17 items as “Sufficiently Qualified and 24 items as “weak.” This indicates that some items are already capable of effectively distinguishing students’ abilities but there are still many items categorizes as “weak’ that need to be revised to improve the quality of the test instrument.

4. Reliability

Reliability testing is conducted to determine the level of consistency of the test instruments used in the study. Reliability indicates the extent to which an instrument can

provide consistent measurement results when used repeatedly on the same subjects under similar conditions. This is in line with Iqbal & Susanto (2025) opinion that reliability testing aims to ensure internal consistency in research instruments. Thus, a reliable instrument will produce stable and trustworthy data. The reliability level of the instrument was interpreted based on the Cronbach's Alpha value presented below.

Table III.7
Level of Reliability Based on Alpha Value

Cronbach Alpha Value	Reliability level
0.0-0.20	Less Reliable
>0.20-0.40	Rather reliable
>0.40-0.60	Pretty reliable
>0.60-0.80	Reliable
>0.80-1.00	Very reliable

Source (Wahyudi, 2016)

Table III. 8
The Result of Reliability Test

Reliability Statistics	
Cronbach's Alpha	N of Items
.899	28

Reliability test was conducted on 28 items using Cronbach's Alpha formula with the assistance of SPSS 27 software. Based on the test results, a Cronbach's Alpha value of 0.899 was obtained, indicating a "Very reliable". This indicates that the research instrument has a very high level of consistency, making it suitable for use as a data collection tool in this study.

5. Descriptive Statistics

Descriptive statistics were used to describe the characteristics of the data obtained from the pre-test and post-test scores of the experimental and control classes. This analysis provides an overview of students' learning outcomes before conducting further statistical analysis. The descriptive statistics calculated in this study included the number of participants (N), total score (Sum), minimum score (Min), maximum score (Max), and mean score (Mean).

6. Assumption Test

a. Normality test

A normality test was conducted to determine whether the research data was normally distributed or not (Ningrum et al., 2024). This test is important as a prerequisite before using parametric statistical analysis in quantitative research. The data tested for normality in this

study were the pre-test and post-test scores in the experimental and control classes. Then there are testing criteria in the Normality Test, namely:

- 1) The data are normally distributed if the value (Sig.) is greater than 0.05.
- 2) The data are not normally distributed if the value (Sig.) is less than 0.05.

b. Homogeneity Test

The uniformity of variance between groups is an important aspect in experimental research comparing two different groups (Ningrum et al., 2024). Therefore, a homogeneity test was conducted to ensure that the data from the experimental group and the control group had comparable levels of variation. This test aimed to see whether the differences in results were caused by the treatment given, rather than differences in data characteristics between groups. Use the following principles to determine homogeneity:

- 1) The group variances may be homogeneous if the Sig. value is greater than 0.05.
- 2) The group variances may not be homogeneous if the Sig. value is less than 0.05.

7. Hypothesis testing

The parametric analysis in this research using independent sample test test. The t-test is used to determine

whether there is a difference between the mean scores of the two groups being compared. In this study, the *t*-test was applied to compare the pretest and posttest results of vocabulary mastery between the experimental class using the Assemblr Edu application and the control class using conventional learning methods. Because the two groups were independent, the Independent Sample *t*-Test was used. Decisions in the *t*-test are based on significance values.

- 1) If the *p*-value is greater than 0.05 ($p > 0.05$), the null hypothesis (H_0) is accepted and the alternative hypothesis (H_a) is rejected.
- 2) If the *p*-value is less than 0.05 ($p < 0.05$), the null hypothesis (H_0) is rejected and the alternative hypothesis (H_a) is accepted.

