

CHAPTER III

RESEARCH METHOD

A. Research Design

To examine the use of YouTube videos as an aid for brainstorming in English teaching and learning. This study uses a quasi-experimental design with a nonequivalent control group. Quasi-experimental research is a type of quantitative research design used to determine causal relationships between variables when random assignment is not feasible. Unlike true experiments, quasi-experiments lack full randomization of participants but still involve manipulation of an independent variable and measurement of its effect on a dependent variable Creswell (2012).

Quasi-experiments are often employed in real-world settings such as classrooms, where it is difficult or unethical to randomly assign students to different groups. Despite the lack of randomization, the design can still yield meaningful and valid results when executed carefully. In educational research, quasi-experimental designs are widely accepted and used to evaluate instructional methods, curricula, or educational technologies.

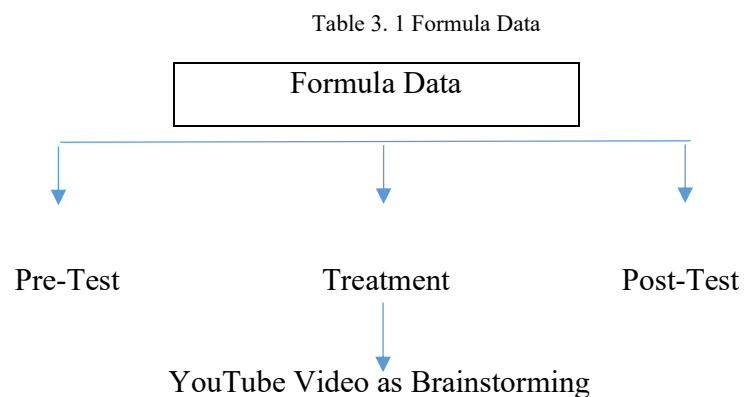
As Highlighted by Gay et al. (2012), this method is appropriate when researchers want to measure the effects of a treatment but cannot randomly assign students due to institutional or ethical constraints. Quasi-experimental

designs are particularly useful in educational contexts, where intact classrooms can serve as treatment and control groups.

Due to logistical constraints, the participants were taken from one intact class consisting of 30 students. The class was divided into two groups: 15 students in the treatment group and 15 students in the control group. The treatment group was taught using YouTube videos as a brainstorming medium, while the control group received traditional instruction.

Both groups were given a pre-test and a post-test to measure their vocabulary improvement. This design is appropriate for school-based research where full randomization is not feasible but valid comparisons between groups can still be made Creswell (2012).

The formula is presented as follows:



The researchers chose research with total population sampling, The researcher chose total population sampling because of the limitations of the new school, so that each academic class only has 1 class, grades 7, 8, and 9 each have 1 class. To support that, reporting from Saunders et al. (2023) mentioning “Total population sampling is a type of purposive sampling technique in which the entire population that meets the criteria is included in the study”. This is also clarified by Sugiyono (2013) “If the population in a study is less than 100, then it is better to take it all, so that the research is population research.” In the case of research, some researchers are sometimes constrained in data collection which sometimes does not have a large population, as highlighted in Creswell (2012) “In some research settings, especially in small-scale studies, the entire population is used as the sample, especially when the size is manageable”.

Given the small size of the population (only one class with 30 students), the entire population was used as the sample in this study. This approach aligns with total population sampling, which is appropriate when the population is limited and accessible. Consequently, all members of the population also served as the participants in the research process. Researchers obtained this sample, comprising a population of 30 individuals, by implementing the theory of De Winter. Contrary to prevailing conventional assumptions, De Winter (2013) demonstrated through extensive simulation studies that the student’s t-test remains a viable inferential method even under such stringent sample size limitations. The simulations

demonstrated that the Type I error rate (false positives) persists at a proximity to the nominal 5% level in the majority of scenarios when samples are obtained from a normally distributed population. Achieving acceptable statistical power (i.e., $\geq 80\%$) is contingent upon the presence of a substantial effect size. Several researchers have noted that quasi-experimental designs are suitable even with small sample sizes as long as careful control and statistical analysis are applied.

As highlighted by De Winter (2013) that inferential statistics, such as the t-test, are still applicable even with small sample sizes (e.g., $n = 15$ per group), as long as data are normally distributed and effect sizes are substantial. As emphasized by De Winter (2013). “Student's t-test remains robust with small sample sizes, provided the assumptions of normality are met”. In addition, Creswell (2012) also asserts that quasi-experiments can be conducted with smaller participant pools when randomization is not possible, and control groups are logically comparable. “In quasi-experiments with 15–30 participants per group, meaningful effect size analysis can still be achieved”. Thus, in studies with 30 students divided into two groups of 15 (treatment and control), it is methodologically acceptable to apply quasi-experimental designs.

B. Research Setting

This investigation took place at SMP Hidayatullah Junior High School in Martapura, South Kalimantan, Indonesia. SMP Hidayatullah Martapura starting in the 2024-2025 academic year implements 2 types of curriculum, namely the

Merdeka Curriculum and the 2013 Curriculum. The Merdeka Curriculum will be implemented in grade 7, while the 2013 Curriculum is still being implemented in grade 8 and grade 9. The curriculum content in the implementation of the Merdeka Curriculum includes subjects: PABP, Civics, Indonesian, English, Mathematics, Science, Social Studies, PJOK, Informatics, Fine Arts and local content (BTA). Meanwhile, the 2013 Curriculum includes: PAI, Civics, Indonesian, English, Mathematics, Science, Social Studies, Physical Education, Crafts, SBK and local content (BTA). The implementation of the learning process at Hidayatullah Martapura Junior High School is carried out in the form of regular learning activities. Regular learning is a learning process that is carried out in class regularly every day from Monday to Saturday. The curriculum content at Hidayatullah Martapura Junior High School contains several components including intracurricular learning content, co-curricular (Pancasila Student Profile Strengthening Project), extracurricular (talents and interests), and school habituation. The Pancasila Student Profile Strengthening Project is implemented in the Merdeka Curriculum (*KSP SMP HIDAYATULLAH MTP*, n.d.). With additional information, researchers also saw that English is also emphasized as a basic medium of learning for students. The decision to conduct research in this school came from the researcher's interest in measuring the extent of students' improvement in the English vocabulary acquisition class where the teacher presented YouTube as a learning medium as a pre-learning/brainstorming. First, the researcher obtained permission from the school by providing them with a valid

research letter. Furthermore, the researcher would conduct the study with students from the selected school.

C. Population and Sample/Participants

The participants in this study included students in the 8th grade class of SMP Hidayatullah Martapura, South Kalimantan, Indonesia. The researcher chose this class because the researcher wanted to focus on the vocabulary test subject where the lesson was implemented in the school. The researcher find that the school has already implement learning school using media technology, that include one of the technologies is YouTube. The total number of students is 30. There are 15 male and 15 female students. Hidayatullah Junior High School, formerly known as Taman Hudaya Pesantren, is a recently established educational institution under the Hidayatullah Foundation. This development follows the establishment of another institution, Hidayatullah Islamic Boarding School, which falls under the purview of the Ministry of Religious Affairs. Consequently, Hidayatullah Junior High School was established under the oversight of the Ministry of Education and Culture.

D. Data

The data source was the primary objective of the research. The data for this study was collected from eighth-grade students from SMP Hidayatullah Martapura Junior High School, who volunteered to participate. In this study, the population size was the participants themselves because, in the case of this new school, there

was only one 8th-grade class, so the class was divided into two groups. During the treatment, the class was separated into two different locations to prevent the teaching materials from being copied by the other group. Additionally, each group received a treatment, with three treatments per group (total of six treatments). In the treatment material, namely: the -ing form of verbs, prepositions of place, and descriptions of people. The same material was taught to both groups but with different approaches. The treatment group was treated using YouTube as a brainstorming medium before learning to enhance students' vocabulary, while the control group was only given traditional brainstorming before learning to enhance students' vocabulary. Data was analyzed using SPSS 26 with a quasi-experimental design, as the study compared two groups and aimed to provide more complex answers. The results of the pre-test and post-test for the treatment group were also calculated. N-Gain was also used in this study to support the data.

E. Technique of Data Collection

The main instrument in this study, a vocabulary test was administered in the form of a written test with 20 questions. This tool was used in pre-tests and post-tests to assess students' vocabulary mastery. before and after the treatment was given.

The types of questions were compiled based on vocabulary mastery indicators and adjusted to the abilities of eighth-grade students. The questions consisted of:

1. 10 multiple choice questions

- a. Each question is worth 4 points for a correct answer.
- b. Used to measure word recognition, synonyms, and usage in context.

2. 4 jumbled sentences

- a. Each question is worth 7.5 points for the correct sentence structure.
- b. Used to measure understanding of sentence structure and accurate word placement.

3. 6 questions on filling in the blanks in the text

- a. Each question is worth 5 points for a correct answer.
- b. Used to assess students' ability to fill in vocabulary based on the context of the reading passage.

The maximum total score for all instruments is 100 points, which is the sum of the weights of all questions.

To ensure the validity and reliability of the collected data, the test was administered following a structured procedure:

1. Pre-Test

Before initiating the conceptualization of classes utilizing the YouTube channel, the researcher administered a pre-test to gauge students' baseline knowledge. The pre-test in this study involved a retelling task, where students were asked to recount the content of the video they had viewed. This assessment was followed by the teacher's utilization of vocabulary relevant to the students' needs.

2. Treatment

In this method, the instructor provide basic vocabulary content based on student needs, which is used in pre-learning as a brainstorming medium. The researcher then assess whether YouTube, as a brainstorming media, can truly boost student vocabulary as needed. In the treatment, there were also three sessions prepared for both groups, namely the first material on verb-ing, the second material on prepositions of place, and the third material on describing people. The treatment was carried out at the beginning before the main learning activity, which was brainstorming. The treatment group was given material on YouTube as the medium, while the control group only used traditional brainstorming activities.

3. Post-Test

In this study, a vocabulary test was administered in the form of a written test with 20 questions. This tool was used in pre-tests and post-tests to assess students' vocabulary mastery. as a brainstorming medium during treatment in the treatment group and those who were only treated using conventional learning methods in the control group.

F. Research Instrument

The researcher collected data for this study employing a vocabulary test as the major research instrument. The test was given as a pre-test and a post-test to assess students' vocabulary mastery before and after the treatment. The instrument

was developed by the researcher and validated through expert judgment by an English teacher.

The test consisted of 20 items, divided into three types of questions to assess different aspects of vocabulary knowledge:

- a. **10 multiple-choice questions**, each with four options, designed to assess students' recognition and understanding of word meanings.
- b. **4 sentence arrangement questions**, requiring students to reorder jumbled words into correct sentences, to test their syntactic awareness and vocabulary usage.
- c. **6 fill-in-the-blank (cloze test) items**, where students completed short texts by inserting appropriate vocabulary words into blanks, aimed at measuring contextual understanding.

Each item type had its own scoring criteria:

- a. Multiple-choice: each correct answer = **4 points**
- b. Sentence arrangement: each correct sentence = **7.5 points**
- c. Fill-in-the-blank: each correct word = **5 points**

The total possible score was **100 points**, and the same structure was used for both pre-test and post-test, although different question sets were used to prevent memorization bias. The test was designed to align with vocabulary levels suitable for eighth-grade students based on CEFR A2-B1 descriptors.

G. Data Analysis

All the data would be collected by Pre-Test and Post-Test, with following steps of procedures such as:

1. Scoring the students answer

The scoring system in this study was designed by considering the complexity and cognitive demand of each test item. Quoted to (Brown, 2004), different item types in a language test require distinct levels of cognitive processing, and therefore, it is reasonable to apply different scoring weights. Similarly, Hughes (2003) emphasizes that weighting items based on their importance and difficulty enhances the overall fairness and construct validity of the test.

Since there are 20 questions, divided as follows: 10 multiple-choice questions, 4 sentence arrangement questions, and 6 questions on filling in the blanks in a text. In accordance with these principles, this study uses the following assessment scheme:

- a. Multiple-choice items: 4 points per correct answer
- b. Jumbled sentence items: 7.5 points per correct answer
- c. Fill-in-the-blank items: 5 points per correct answer

This weighting system allows the test to capture both receptive (recognition-based) and productive (usage-based) vocabulary skills

proportionally. The total score possible is 100 points, with clear criteria that reflect the varied nature of vocabulary knowledge being assessed. Moreover, to avoid test-retest bias and memory recall effects, the pre-test and post-test were constructed using different but equivalent questions. This follows Creswell (2012) recommendation that pre- and post-assessments should be matched in structure and level of difficulty without using identical items.

2. Classifying the score levels classification is a follow:

Table 3. 2 Classification student's score:

NO.	Classification	Score
1.	Excellent	86-100
2.	Good	71-85
3.	Fair	56-70
4.	Poor	41-55
5.	Very Poor	< 40

Following the collection of data from students through pretests and posttests, the researcher conducted an analysis of the results. The pre-test and post-test scores were calculated, and the total number of correct answers was determined. The vocabulary test results were then tabulated using a T-test, and statistical analysis was performed using SPSS (Statistical Package for Social Sciences) to examine

whether there was a statistically significant effect on students' gain scores Puskur (2006).

Given the single-group design used in this investigation, a normality test was required to evaluate if the pre-test and post-test data were normally distributed. The Shapiro-Wilk test, utilizing SPSS 26, was applied to assess the normality of the data. A significance level of 0.005 was employed, indicating that if the test result exceeded this threshold, the data would be deemed normally distributed.

In addition to statistical tests such as independent samples t-test and paired samples t-test, the researcher also used Normalized Gain (N-Gain) to evaluate the effectiveness of the treatment. The N-Gain was calculated to determine the degree of improvement in students' vocabulary mastery. This method, proposed by Hake (1999), allows for a more accurate interpretation of learning gains, especially in quasi-experimental designs.

The formula used to calculate the N-Gain is:

Table 3. 3 Formula N-Gain

$$\text{N-Gain} = \frac{(\text{Post-Test Score}) - (\text{Pre-Test Score})}{(\text{Maximum Score} - (\text{Pre-Test Score}))}$$

According to Hake (1999), the criteria for interpreting N-Gain values are as follows:

- a. High: ≥ 0.70
- b. Moderate: $0.30 - 0.69$
- c. Low: < 0.30

This analysis was specifically applied to the treatment group to determine the level of learning gain achieved through the YouTube-based brainstorming strategy.

H. Technique of Data Validity

Verification is importance to measure of validity test. To ensure the content validity of the vocabulary test used in this study, the instrument was reviewed by an English teacher with relevant experience teaching eighth-grade students. The expert was asked to evaluate each test item in terms of clarity, relevance to the learning objectives, difficulty level, and appropriateness for the students' level.

This expert judgment process follows the guideline of content validity as explained by Suharsimi (2010), who states that expert review is an acceptable method to establish the validity of test instruments in educational research. The feedback provided by the expert was used to revise and refine the test items before they were administered to the students.