

CHAPTER III

RESEARCH METHOD

A. Research Design

This study uses a quantitative method, which is a type of research that emphasizes the in-depth collection of data and covers a broad population scope. Quantitative research is characterized by being systematic, planned, and well-structured from the initial stage through the design of the study. The approach used in this study is experimental, one of the forms of quantitative research that is very effective in measuring cause-and-effect relationships. This study falls into the category of field experiments, where the group receiving the treatment or stimulation is not separated from their everyday environment. This approach has the added advantage of allowing the researcher to observe other independent variables that may influence changes in attitudes (Listiorini 2022)

Quasi-experimental research is an experimental approach used when full control over all variables is difficult especially in social research settings like classrooms (Abraham & Supriyati, 2022). This experimental design is a development of the quasi experimental design, which is often difficult to implement. Although it has a control group, this design is not fully capable of controlling all additional variables that could affect the experimental results. However,

this design is considered better than the pre-experimental design. Quasi-experimental designs are often used because, in practice, it is difficult to obtain an ideal control group for research. In the context of learning, for example, it is often not possible to separate students into experimental and control groups. For instance, some students may learn using the discussion method, while others do not. To overcome the difficulty of determining a control group in such research, the quasi-experimental design was developed (Fenti Kimawati, 2017).

In this design, both groups are given a post-test (O) after the treatment is applied. The procedure involves dividing the subjects into two groups, where the experimental group receives a specific treatment, while the control group does not receive any treatment. The design used is the Nonequivalent Posttest-Only Control Group Design, which allows for the evaluation of the treatment outcomes in both groups. The design can be illustrated as follows:

Table 2. Research Design

Group	Pre-test (O1)	Treatment (X)	Post-test (O2)
Experimental	O1	X	O2
Control	O3	O	O4

Source:

(Sugiyono, 2019)

Description:

- O1 = Pre-Test Value of Experiment Class
- O2 = Post-Test Value of Experiment Class
- O3 = Pre-Test Value of Control Class
- O4 = Post-Test Value of Control Class
- X = Treatment (Learning Using Tiktok Application)
- O = Not applying the Reverse Class Model
- O2 = Post-Test Value (After Treatment)

This experimental model consists of two main steps. First, the experimental group receives treatment through the use of the TikTok application, while the control class is taught without incorporating TikTok, and then a post-test is conducted to measure students' vocabulary mastery after the treatment, with the aim of identifying whether there is a significant after using a TikTok application.

B. Location and Time of Research

This research was conducted at SMPN 05 Bengkulu City. Based on initial observations on September 12, 2024 to October 30, 2024. The Research will conduct on september to october 2025. The subjects of this study were students of class VIII G which amounted to 33 students.

C. Population and Sample

1. Population

In research, a population is defined as a group of objects or subjects that possess certain traits and characteristics, which are selected by the researcher to be

studied and later concluded. A population can be understood as the entirety of individuals who become the main focus of the study and will subsequently undergo the process of generalization. Generalization itself is a way to draw conclusions about a broader group based on data obtained from a smaller portion of the population. Therefore, the researcher determines that the population includes all objects or subjects being studied, whose results will be generalized in order to formulate conclusions from the collected data. In this quasi-experimental study, the population consisted of all eighth-grade students of SMPN 05 Bengkulu City in the academic year 2024/2025.

Table 3. Research Population

No	Class	Female	Male	Total
1.	8 A	17	18	35
2.	8 B	18	18	36
3.	8 C	22	14	36
4.	8 D	21	15	36
5.	8 E	22	11	33
6.	8 F	17	19	36
7.	8 G	20	13	33
8.	8 H	20	16	36
9.	8 I	18	18	36
10.	8 J	16	19	35
11.	8 K	16	20	36
Total				391

2. Sample

The sampling technique used in this study was purposive sampling. According to Sugiyono (2017),

purposive sampling is a technique for selecting a sample based on specific criteria determined by the researcher. The selected class was considered to have characteristics that aligned with the research objectives and represented the population under study. This technique was chosen because it allows the researcher to select a sample that best fits the needs of the study. The criteria used included the similarity of students' characteristics, learning environment, and curriculum implementation. Based on these considerations, class VIII G, consisting of 33 students, and VIII E as control class consisting of 33 students was selected as the sample. This design was considered appropriate for examining the effect of the treatment while controlling individual differences among participants. Therefore, the changes in students' vocabulary achievement could be attributed more directly to the treatment provided.

The selection of the sample is based on specific considerations to ensure that the participants represent the characteristics of the population. According to Sugiyono (2017), sampling techniques play an important role in determining the validity and reliability of research results. Therefore, the researcher ensures that the selected sample is appropriate and relevant to the objectives of the study.

Table 4. Research Sample

No	Class	Classroom	Female	Male	Total
1.	Experimental Class	8 G	20	13	33
2.	Control Class	8 E	22	11	33
Total			42	24	66

D. Definition of Variabels

This research involves two types of variables: the dependent variable, which is affected by other variables, and the independent variable, which impacts other variables. The variables in this research are as follows:

1. TikTok Application, the use of English video content on TikTok by students is the activity of watching, understanding, and utilizing English videos uploaded on the TikTok platform. The definition of English video content is any form of video that uses English as a medium of communication, which may include various genres such as tutorials, entertainment, education, and daily conversations.
2. Student vocabulary Vocabulary mastery is the ability of students' to understand, remember, and use English vocabulary obtained from video content on Tiktok appropriately and according to context.

E. Data Collection Technique

In this study, pre-test and post-test were used to measure students' vocabulary mastery before and after the treatment.

The use of these two tests aimed to identify differences in students' abilities more clearly. The treatment used English video content on TikTok to provide engaging and authentic vocabulary exposure. This media was chosen because it is considered capable of increasing students' learning interest. The test, in the form of multiple-choice questions, was based on the taught vocabulary material. The items were designed to measure students' understanding objectively. The post-test was then administered to evaluate improvement in vocabulary mastery and the effectiveness of TikTok-based learning. Thus, the results obtained are expected to show the effect of the given treatment. The steps of data collection are as follows:

a. Pre test

The pre-test was administered to assess the students' English language proficiency, particularly in the area of vocabulary. The pre-test was conducted prior to the intervention to determine whether the intervention had any effect before administering the post-test.

b. Treatment

The treatment in this study are learning English vocabulary through the TikTok application, In the experimental group, the use of TikTok was integrated into the teaching and learning process as a medium for delivering vocabulary materials. Before each meeting, the researcher prepare and share short TikTok videos

containing target vocabulary items, pronunciation guides, example sentences, and visual or contextual illustrations. These videos were designed to be concise, interesting, and relevant to the students' daily lives. During the treatment sessions, students were encouraged to watch the videos either before class or at the beginning of the lesson, then participate in follow-up activities such as vocabulary matching games, pronunciation practice, sentence construction, and short discussions based on the video content. The teacher facilitated the learning process, guided students in using the vocabulary in context, and provided clarification when necessary.

The integration of TikTok aimed to make vocabulary learning more engaging and meaningful by combining visual, auditory, and contextual elements. Through the short video format, students were expected to gain better retention of words and greater motivation to learn English. This method also promoted learner autonomy, as students could rewatch the videos anytime for review. However, the use of TikTok also presented certain challenges, such as dependency on internet connection, the potential for distraction from non-educational content, and limited explanation depth compared to traditional materials.

The treatment will be conducted in several meetings over 6 sessions with several learning topics to be taught, such as nouns, verbs, adjectives, synonyms, and antonyms. Each meeting lasted 70 minutes, and students were given worksheets to assess their learning progress per meeting. The purpose of this treatment was to compare the effectiveness of using TikTok as a digital learning tool with traditional book-based methods in improving students' English vocabulary mastery.

Tabel 5. Teaching Treatment for Experimental Class

Aspect	Experimental Class
Learning Method	The experimental group was taught using the TikTok application as the main medium for vocabulary teaching. TikTok was integrated into the learning process to present vocabulary in a more interesting and contextual way. The researchers prepared short educational videos for each session containing target words, pronunciation guides, example sentences, and visual illustrations related to real-life situations. These videos were designed to attract students' attention and make vocabulary learning more meaningful by combining visual and auditory input.
Learning Activities	Students in the experimental group were asked to watch TikTok videos before class or at the beginning of each lesson. These videos formed the basis for classroom activities, such as vocabulary matching games, pronunciation exercises, sentence construction, and short discussions. Through these activities, students were encouraged to actively use new vocabulary in meaningful contexts, rather than simply memorizing lists of words. Teachers acted as facilitators, guiding students in the correct use of

	vocabulary and providing explanations when necessary.
Learning Atmosphere	The integration of TikTok creates a dynamic and interactive classroom environment. Students are generally more enthusiastic and motivated because the videos present vocabulary in a fun and familiar format. The combination of visual, auditory, and contextual elements helps students understand and remember new words more effectively. In addition, because the videos are accessible at any time, students are encouraged to review the material independently, which promotes greater independence and responsibility in their learning process.
Durations and Topics	Treatment in the experimental group was conducted over six sessions, each lasting 70 minutes. Each session focused on a specific vocabulary topic, such as nouns, verbs, adjectives, synonyms, and antonyms. At the end of each session, students completed worksheets designed to assess their understanding and application of the new vocabulary learned through TikTok videos.

c. Post-test

After the completion of the treatment, a post-test was administered to determine the impact of TikTok video content on the students' vocabulary mastery. This test, similar in format to the pre-test, measured the students' ability to recognize, understand, and use English vocabulary. The primary objective of the post-test was to assess whether there was a significant improvement in vocabulary scores compared to the pre-test. The results were recorded, analyzed, and interpreted to evaluate the effectiveness of the treatment. The classification of student

scores in both pre-test and post-test is based on Depdikbud (1985) as follows:

Table 6. Vocabulary Assessment Rubric

Score	Classification	Contextual Understanding	Lexical Recognition
96-100	Excellent	Accurately understands and applies a wide range of vocabulary (nouns, adjectives, synonyms, antonyms, and word classification) in various sentence contexts with minimal errors	Demonstrates strong mastery of word meanings, spelling, and usage; consistently recognizes and applies vocabulary correctly.
86-95	Very Good	Understands most vocabulary in context with only minor interpretation errors.	Recognizes and uses vocabulary accurately in most situations; few mistakes that do not hinder meaning.
76-85	Good	Understands common vocabulary well, though sometimes confused by less familiar words or abstract meanings.	Generally identifies correct word meanings but may misuse some vocabulary in less familiar contexts.
66-75	Fairly Good	Understands basic vocabulary with some effort; context sometimes helps clarify meaning.	Recognizes simple words correctly but struggles with complex or similar-looking ones.
56-65	Fairly	Often misinterprets meanings or relies on guessing when context is unclear.	Limited recognition of vocabulary; frequent confusion between meanings and spelling.
36-55	Poor	Rarely understands vocabulary in context; depends heavily on guessing.	Weak vocabulary knowledge; unable to distinguish correct word meanings.
0-35	Very Poor	Fails to understand vocabulary in context;	Shows almost no understanding of vocabulary meaning or

		responses are mostly inaccurate or irrelevant.	usage; answers mostly random or incorrect.
--	--	--	--

d. Documentation

The process of giving users organized information about a machine learning model so they may comprehend its development, performance, limitations, and proper application is known as documentation. By giving precise information about the model's architecture, training data, evaluation outcomes, and planned applications, it helps to increase transparency, improve interpretability, and facilitate well-informed decision-making. Additionally, documentation ensures responsible and ethical use, helps users identify the model's strengths and limitations, identifies potential dangers or biases, and promotes efficient communication between developers, researchers, and other stakeholders (Crisan, Drouhard, and Vig n.d.).

F. Research Instrument

The research instrument is a tool used by researchers to collect data in order to answer research questions. An appropriate instrument is very important to obtain accurate and reliable data. In this study, the instrument used is a vocabulary test in the form of multiple-choice questions. This test is designed to measure students' vocabulary mastery before and after the treatment. Each correct answer is scored 1, while incorrect answers are scored 0 on a scale of 0–1. This scoring

system aims to provide objective and measurable results. The final scores are then converted to represent the level of students' vocabulary mastery. To ensure validity, the test items are reviewed by English lecturers, and additionally, the validation process uses statistical tests in the form of validity and reliability tests. The validation process emphasizes content relevance, clarity of language, and appropriateness of the difficulty level, so that the test can accurately measure students' vocabulary mastery. Thus, the instrument used is expected to produce valid and reliable data.

The instrument used in this study is a vocabulary test designed to measure students' vocabulary mastery before and after the treatment. The test includes aspects such as meaning, spelling, pronunciation, and word usage in context. To ensure the quality of the instrument, validity and reliability tests are conducted. According to Lyle Bachman, a good instrument must be both valid and reliable in order to accurately measure students' language ability.

G. Data Analysis Technique

The researcher analyzed the data using the pre-test and post-test results from the experimental group. The goal is to see if the application of this Tiktok Application can affect the level of students' understanding in English skills, especially in Vocabulary Mastery (Nurhidayah 2020).

1. Validity Test

Following (Listiorini 2022) validity measures how well an instrument captures what it wants to assess. If the significance value < 0.05 , the item is considered valid. The score of each question will be correlated with the total score. An item is valid if the $r\text{-count} > \text{the } r\text{-table}$; if the $r\text{-count} \leq \text{the } r\text{-table}$, the item is considered invalid. Overall the data computing out with aid of computer facilities SPSS ver 26.

In this validity test, the researcher used multiple choice question, using a total 50 questions. The researcher conducted a trial on students who has a same grade as students that researcher will use in research later, the trial was using 15 students as the subjects of trial. After carrying out the trial, the researcher processed the data of the questions by calculating the validity of each questions. Therefore, the researcher found total 23 questions valid out of 50 questions. To determine the validity item, each item $r\text{-count}$ was compared to the $t\text{-table}$ value, which is 0.514 at a significance level 5% for $N=15$. The instrument was analyzed using SPSS 26, Based on the decision rule, if $r\text{-count} > 0.514$ and $sig < 0.05$, otherwise if the $r\text{-count} < 0.514$ the items considered invalid.

Tabel 7. Indicators of Vocabulary Test Validity

No	Indicator	Number of Valid Items	Valid Item Numbers
1	Nouns	5	1, 4, 11, 21, 26
2	Verbs	6	2, 7, 12, 24, 27,32
3	Adjectives	6	8, 14, 18, 23, 28, 44
4	Synonyms	3	5, 15, 25
5	Antonyms	3	10, 20, 30
Total		23	

Based on the results of the validity analysis that has been conducted, and based on the criteria below,

No	R-Count	Validity Level
1.	0.00-0.49	Very Low
2.	0.50-0.69	Low
3.	0.70-0.79	High
4.	0.80-1.00	Very High

(Source: Widodo, et.al. 2019)

Based on the results of the criterion validity analysis and the validity data analysis for this research instrument, the results of the validity test are presented in the following table

Tabel 8. The Instrument Validity

Item	r Table	Significance	r Count	Results
1	0.514	0.000	0.868	Valid
2	0.514	0.000	0.834	Valid
3	0.514	0.348	0.261	Unvalid
4	0.514	0.045	0.523	Valid
5	0.514	0.004	0.690	Valid

6	0.514	0.493	0.192	Unvalid
7	0.514	0.008	0.690	Valid
8	0.514	0.019	0.596	Valid
9	0.514	0.450	0.211	Unvalid
10	0.514	0.028	0.564	Valid
11	0.514	0.000	0.801	Valid
12	0.514	0.000	0.834	Valid
13	0.514	0.662	0.123	Unvalid
14	0.514	0.001	0.753	Valid
15	0.514	0.000	0.800	Valid
16	0.514	0.066	0.486	Unvalid
17	0.514	0.101	0.440	Unvalid
18	0.514	0.022	0.585	Valid
19	0.514	0.167	0.376	Unvalid
20	0.514	0.012	0.627	Valid
21	0.514	0.000	0.801	Valid
22	0.514	0.007	0.662	Valid
23	0.514	0.032	0.554	Valid
24	0.514	0.003	0.717	Valid
25	0.514	0.038	0.538	Valid
26	0.514	0.006	0.677	Valid
27	0.514	0.000	0.878	Valid
28	0.514	0.001	0.782	Valid
29	0.514	0.057	0.502	Unvalid
30	0.514	0.039	0.538	Valid
31	0.514	0.139	0.400	Unvalid
32	0.514	0.380	0.244	Unvalid
33	0.514	0.689	0.113	Unvalid
34	0.514	1.000	0.000	Unvalid
35	0.514	0,519	-0,181	Unvalid
36	0.514	0,264	0,308	Unvalid
37	0.514	0,622	0,139	Unvalid
38	0.514	0,066	-0,486	Unvalid
39	0.514	0,902	-0,035	Unvalid
40	0.514	0,573	0,158	Unvalid
41	0.514	0,913	-0,031	Unvalid
42	0.514	0,185	0,362	Unvalid

43	0.514	0,418	0,226	Unvalid
44	0.514	0,012	0.627	Valid
45	0.514	0,827	-0,062	Unvalid
46	0.514	0,666	-0,122	Unvalid
47	0.514	0,456	0,208	Unvalid
48	0.514	0,686	-0,114	Unvalid
49	0.514	0,150	0,391	Unvalid
50	0.514	0,743	-0,092	Unvalid

In this validity test, the researcher used 50 multiple-choice questions that were developed based on predetermined indicators. The instrument was piloted at Pondok Pesantren Salafiyah Hidayatul Qomariyah, Bengkulu City, which has an accreditation level equivalent to the main research site, namely SMP Negeri 5 Bengkulu City. The pilot test involved 17 students as the participants to evaluate the validity of the test instrument. After the pilot test was completed, the researcher analyzed the data by calculating the validity of each test item. Based on the results of the analysis, it was found that 23 out of the 50 test items were valid according to the predetermined indicators:

Table 9. Indicators of Vocabulary Validity Test

No	Test item	Test indicator
1.	Which object is used to write? a. Spoon b. Pen	Students are able to identify the English noun for a writing tool.

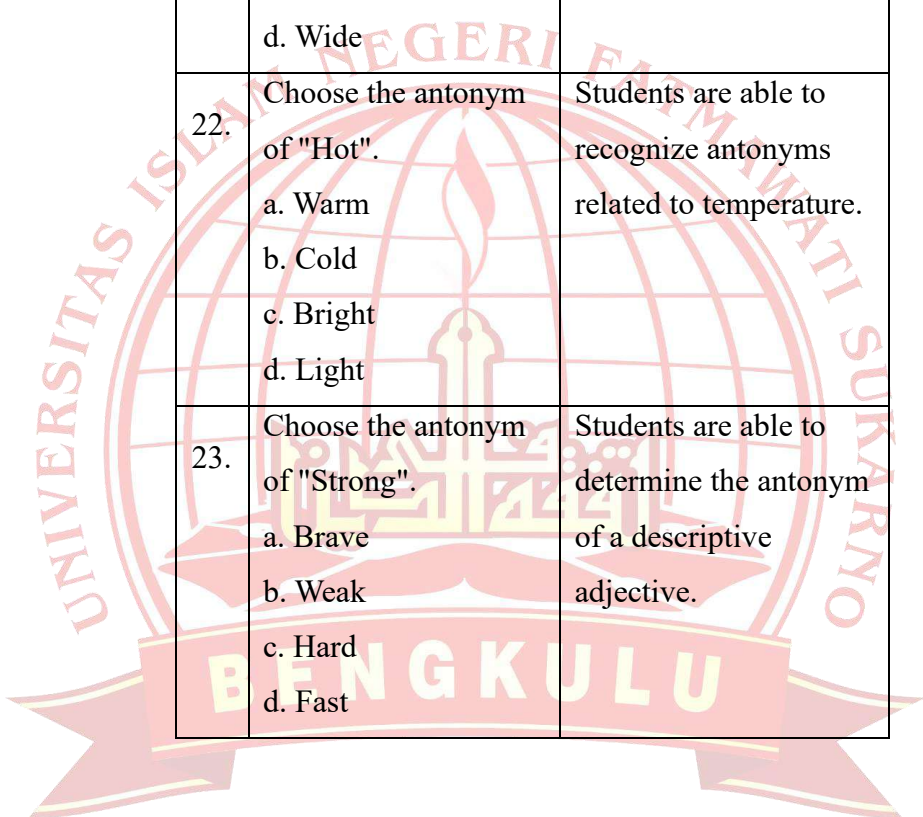
	c. Plate d. Chair	
2.	Which word is a noun? a. Run b. Chair c. Eat d. Sleep	Students are able to distinguish nouns from verbs.
3.	Which object is used to cut paper? a. Scissors b. Brush c. Ruler d. Key	Students are able to recognize the English vocabulary for classroom objects.
4.	Which object is used to open a locked door? a. Bottle b. Key c. Bag d. Spoon	Students are able to identify nouns related to everyday objects.
5.	Which object is used for drinking water? a. Plate	Students are able to identify nouns based on their functions.

	b. Glass c. Pen d. Chair	
6.	Which verb means to eat food? a. Drink b. Run c. Eat d. Sleep	Students are able to identify the meaning of common action verbs.
7.	Which verb means to move quickly with your legs? a. Run b. Sit c. Read d. Draw	Students are able to recognize verbs describing physical movement.
8.	Which verb means to go to bed and rest? a. Jump b. Sleep c. Swim d. Write	Students are able to identify verbs related to daily activities.
9.	Which word is an action verb? a. Teacher	Students are able to classify action verbs correctly.

	b. Run c. Bag d. Chair	
10.	Which verb means to see with your eyes? a. Run b. Watch c. Close d. Read	Students are able to identify verbs related to the sense of sight.
11.	Which verb means to use your ears to understand sound? a. Listen b. Jump c. Cook d. Play	Students are able to recognize verbs related to the sense of hearing.
12.	Which adjective describes something that has a lot of light? a. Dark b. Bright c. Weak d. Cold	Students are able to identify adjectives describing physical characteristics.
13.	Which word is an adjective?	Students are able to distinguish adjectives

	a. Tall b. Pencil c. Window d. School	from other parts of speech.
14.	Which adjective describes a person who has a lot of strength? a. Strong b. Weak c. Thin d. Soft	Students are able to identify adjectives describing personal qualities.
15.	Which adjective means "not slow"? a. Fast b. Short c. Lazy d. Old	Students are able to determine the meaning of descriptive adjectives.
16.	Which adjective means "having a low temperature"? a. Hot b. Warm c. Cold d. Bright	Students are able to recognize adjectives related to temperature.

17.	Which word is an adjective? a. Bag b. Big c. Drink d. Door	Students are able to classify adjectives correctly.
18.	Choose the synonym of "Happy". a. Sad b. Angry c. Glad d. Weak	Students are able to identify the synonym of a common adjective.
19.	Choose the synonym of "Fast". a. Quick b. Slow c. Lazy d. Heavy	Students are able to recognize words with similar meanings.
20.	Choose the synonym of "Beautiful". a. Ugly b. Pretty c. Sad d. Tall	Students are able to determine the correct synonym of a given word.



21.	Choose the antonym of "Big". a. Small b. Tall c. Long d. Wide	Students are able to identify the opposite meaning of common adjectives.
22.	Choose the antonym of "Hot". a. Warm b. Cold c. Bright d. Light	Students are able to recognize antonyms related to temperature.
23.	Choose the antonym of "Strong". a. Brave b. Weak c. Hard d. Fast	Students are able to determine the antonym of a descriptive adjective.

2. Reliability Test

Reliability testing is a tool used to evaluate the consistency of a questionnaire, which serves as an indicator of a variable. If a measuring instrument consistently produces stable results, it is considered reliable. Essentially,

reliability measures the consistency of a questionnaire as an indicator of a variable or construct. Various methods can be used for reliability testing, including the Alpha coefficient method. In this approach, scores are divided into halves based on the number of questionnaire items and analyzed using Reliability Analysis. A question item is deemed reliable if it achieves a value of 0.600 or higher, indicating that only 40% of the test scores reflect error variations (Listiorini 2022).

In this reliability test, the researcher used multiple-choice questions with a total of 50 questions. Then the researcher conducted a trial in one of the high schools that had the same grade as the school that the researcher would use for the place where the researcher would conduct research later, the trial was carried out with a total of 15 students as the subject of the trial, after carrying out the trial the researcher processed the data on the test questions by calculating the reliability. After processing the reliability data on the test questions, the coefficient of the reliability test is 0.883 which was explained earlier that the question item is considered reliable if it reaches a value of 0.600 or higher therefore it can be assumed that this test has a high level of reliability.

Table 10. The Result of Reliability

Reliability Statistics	
Cronbach's Alpha	N of Items
0,883	50

Based on the table above, it can be concluded that the 50 questions, tested on a total of 15 students, are highly suitable for use in the pre-test and post-test to be conducted by the researcher. This is supported by the fact that the Cronbach's Alpha value reached 0.883. The value of Cronbach's Alpha can be interpreted as follows, according to Taber (2018):

Table 11. Cronbach's Alpha Interpretation

Cronbach's Alpha	Interpretation
0.11-0.44	Low
0.45-0.65	Acceptable
0.58-0.68	Slightly Low
0.70-0.77	Fairly High
0.81	Robust
0.84-0.90	Reliable
0.91-0.93	Strong
0.93-0.94	Excellent

3. Item Difficulty Test

Suharsimi Arikunto, as cited in Nasikhudin et al. (2024), states that item difficulty testing is conducted to determine how many respondents answered the item correctly. Item difficulty testing is a crucial aspect of research instrument development. This process involved administering a test consisting of 60 items to the students. These questions were given to 30 eleventh-grade students at the school where the researcher tested the instrument.

Item difficulty is calculated by dividing the number of students who answered a specific item correctly by the total number of students who took the test, this calculation uses the following formula:

$$p = \frac{RH}{NH} + \frac{RI}{NI} \times 100\%$$

Description:

P: Difficulty level in percent

Nh: Number of test takers in the highest score group

Rh: Number of correct answers in the highest score group

Ni: Number of test takers in the low score group

Ri: Number of correct answers in the low score group

$$p = \frac{213}{7} + \frac{105}{7} \times 100\%$$

p: 45.4%

Table 12. Interpretation of Item Difficulty Test

Question Category	Interpretation
0%-15%	Very Difficult
16%-30%	Difficult
31%-70%	Medium
71%-85%	Easy
86%-100and	Very Easy

4. Differentiability Test

Test differentiation is a test designed to measure the difference between high-ability students and low-ability students. The figure indicating the magnitude of this ability difference is called the discrimination index. (Solichin, 2017). The process of testing the items' discriminatory power was conducted by administering a test consisting of 60 items to 30 eleventh-grade students at the school where the researcher tested the test instrument. Items with a discrimination index > 0.30 were deemed good, whereas items with a discrimination index < 0.30 were deemed poor. In calculating this discrimination power, the researcher used Microsoft Excel; to obtain the discrimination index in this test, we used the following formula:

$$DP = \frac{BA}{JA} - \frac{BB}{JB} : NM$$

Description:

DP: Question Differentiation Power

BA: Number of correct answers from high score group test takers

BB: Number of correct answers from low score group test takers

JA: Number of high score group test takers

JB: Number of high score group test takers

NM: Scor Maximum

$$DP = \frac{213}{7} - \frac{105}{7} : 50$$

DP: 0,38

Based on the above calculations, it can be concluded that the discriminative power of each item tested on 15 students using a total of 50 items is considered enough according to the interpretation of the test's difficulty level (Solichin, 2017).

Table 13. Interpretation of Item Differentiability Test

Differentiability Test	Interpretation
0,00 -----0,20	Poor
0,21 -----0,40	Enough
0,41 -----0,70	Good
0,71 -----1,00	Very Good
Negative, all of them	Not good, so all items that have a negative discriminating power value should be discarded.

5. Normality Test

The normality test determines whether the collected data follows a normal distribution, ensuring that statistical assumptions are met. This test is an important initial step before conducting further data analysis. SPSS is used to conduct this test to check the normality of the pre-test and post-test data. The use of SPSS helps simplify the data analysis process in a more accurate and systematic way. If the data is normally distributed, parametric statistical tests, such as the T-test, can be used. Conversely, if the data is

not normally distributed, non-parametric tests can be used as an alternative.

6. Homogeneity Test

The homogeneity test examines whether the sample population has the same variance, which is crucial for making accurate comparisons between the experimental and control groups. The following criteria are used for the homogeneity analysis:

- a. Significance value (α) = 0.05
- b. If $\text{sig} > \alpha$, the data is considered homogeneous (same variance).
- c. If $\text{sig} < \alpha$, the data is not homogeneous (different variance).

Over all the data computing out with aid of computer facilities SPSS ver 28.

7. T-Test

The T-test is used to assess the statistical significance of the differences between the pre-test and post-test scores of the experimental and control groups. This test identifies whether collaborative learning has a significant effect on students' integrated language skills. The following criteria are applied for the T-test analysis:

- a. If the significance level $< \alpha$ (0.05), this indicates a significant effect of flipped classroom toward students integrated language skills.

- b. If the significance level $> \alpha$ (0.05), this indicates no significant effect flipped classroom toward students integrated language skills

This test is conducted using SPSS 29.

8. F-Test

The F-Test is used to examine the impact of two independent variables on a single dependent variable simultaneously. This test helps determine if each independent variable significantly affects the dependent variable and whether there is an interaction between the two independent variables. In this research, F-Test is used to measure the effect of using Tiktok Videos in improving students' vocabulary, therefore the factors that will be used in this test are pretest, treatment, and the interaction between the two whether it will affect the value of the parameter being tested, namely the posttest. The following criteria are applied for the F-test:

- a. If F-count is greater than F-table H_0 is rejected,, then indicating a significant effect.
- b. If F-count is less than or equal to F-table, then H_0 fails to be rejected, indicating no significant effect.