

CHAPTER III

RESEARCH METHODOLOGY

This chapter outlines the research strategy and operational steps utilized to investigate the effect of Gemini AI-synthesized structured materials on students' writing skill. The overview provided here covers the underlying methodology, target participants, research instruments, including the writing test and scoring rubric, data collection procedures, and the statistical techniques used for data analysis to verify the robustness and credibility of the findings.

A. Research Design

This study adopts a quantitative methodology featuring a quasi-experimental setup. According to Creswell (2014), experimental designs are used to test the extent to which a treatment affects certain outcomes or variables. This design is chosen because the situation in the field does not allow for random assignment of individuals to subjects, so the research is conducted in classes that are formed naturally in schools.

The specific research setup utilized is a Non-Equivalent Control Group Design. This configuration involves two distinct groups, specifically an experimental class and a control class. Both cohorts undergo a pre-test to evaluate their baseline performance prior to the intervention. Subsequently, the experimental group receives instruction using structured materials synthesized by Gemini AI, whereas the control group undergoes traditional teaching procedures. Following the intervention period, both classes complete a post-test to statistically measure and contrast their rates of score improvement (Creswell, 2014).

B. Population and Sample

The target population for this research comprises eleventh-grade learners at SMA Negeri 1 Ngimbang during the 2025/2026 academic year. The selection of this population is based on an initial survey and preliminary observations which indicated that students are still encounter significant difficulties in organizing ideas and utilizing structured vocabulary in their writing. Furthermore, the institutional framework of SMA Negeri 1 Ngimbang—characterized by its “double-track” system and its designation as a Sekolah Taruna Tahfidzul Quran (STTQ)—demonstrates a dedication to scholarly growth and technological implementation, rendering it an ideal setting for examining how Gemini AI influences the process of english writing.

The research sample consists of two classes with similar characteristics and levels of English proficiency. These participants are chosen through purposive sampling, a selection technique that takes certain considerations into account in order to achieve specific research objectives. Based on discussions with the English teacher at SMA Negeri 1 Ngimbang, one classroom was designated as the experimental group to receive lessons via Gemini AI-synthesized structured materials, whereas a separate classroom was assigned as the control group that received conventional instruction.

Initially, each class consisted of approximately 30 students. However, not all students were included in the final data analysis. Only students who participated in all stages of the study—including the pre-test, treatment sessions, and post-test—were included in the analysis. Students who were

absent during one or more stages of the study were excluded to ensure data completeness and consistency. As a result, the final sample consisted of 20 learners in the experimental group and 15 learners in the control group. Key considerations in selecting these two classes included:

1. Both classes have relatively balanced average writing scores based on previous academic records.
2. Students in both classes have equal access to the technological facilities needed during the research process.
3. Both classes were taught by the same teacher to minimize any teaching bias that might affect the final research results.

C. Research Instruments

In quantitative research, instruments are tools used by researchers to measure variables objectively. Creswell (2012) states that in experimental designs, instruments are used to collect numerical data that allows researchers to test hypotheses and observe statistical differences between groups. The instruments used in this study include:

1. Writing Test

The primary instrument used in this study was a writing test administered as a pre-test and a post-test (Appendix 4). Creswell (2012) emphasizes the importance of using equivalent measurement instruments in experiments to ensure that changes in scores are truly attributable to the treatment administered. This test uses topics tailored to the students so that they can express their ideas more fluently without being hindered by a lack of basic knowledge.

2. Documentation

Documentation serves as a support tool that ensures the validity and transparency of research and promotes a structured research process. In this study, documentation tools consist of several teaching materials and visual evidence:

- a. **Scoring Rubric:** This is a crucial documentation tool used to evaluate the students' writing objectively. Adapted from Weigle (2002), it uses a 1–5 scale across five analytical aspects: Content, Organization, Vocabulary, Language Use, and Mechanics (Appendix 5). This rubric acts as the official record-keeping tool for quantifying the test results.
- b. **Lesson Plan (RPP):** The lesson plans document the research treatment's design (Appendices 6 and 7). They outline the systematic steps of integrating Gemini AI-synthesized materials into the teaching-learning process, ensuring pedagogical consistency.
- c. **Student Worksheet (LKPD):** These worksheets document the students' learning activities during the intervention, containing the structured materials generated by Gemini AI that provide scaffolding for the writing process (Appendix 8).
- d. **Photographs of Research Activities:** Visual documentation is collected to provide evidence of the research's implementation in the classroom. This includes photos of the researcher delivering the treatment. These photos function as a verification of the study's authenticity and practical execution (Appendix 11).

By integrating these tools, researchers can ensure that the data collected is not only accurate in terms of numerical scores but is also supported by concrete evidence of student engagement with Gemini AI-synthesized materials.

D. Data Collection Procedures

The data collection protocols in this investigation are systematically structured to guarantee that the gathered information aligns with the quasi-experimental quantitative framework. According to Creswell (2012), gathering data within experimental inquiries entails organized steps to objectively evaluate the influence of an intervention. This process is partitioned into three primary phases: the pre-experimental stage, the treatment stage, and the post-experimental stage.

1. Pre-Experiment Stage (Pre-Test)

Before the learning process begins, , a pre-test is administered to both the experimental and control groups. Students are required to complete a writing assignment centered around a personalized topic. Then, the data obtained from the pre-test is used to evaluate the students' initial writing abilities and to verify that both classes have an equal starting point before the intervention is given.

2. Treatment Implementation Stage

This stage is the core of the investigation where the independent variable is actively applied. The comparison between the treatments administered in the experimental and control groups is presented (Appendix 3). The differences in procedures between the two classes are as follows:

- a. Experimental Class (Learning with Gemini AI): Students learn writing with the help of Gemini AI-Synthesized Structured Materials (Appendix 8). The procedure in this class follows this flow:
 - i. Introduction Session: The researcher demonstrates how to use worksheets containing structured materials synthesized by Gemini AI. Students are guided to understand how text elements (vocabulary and framework) can help their writing process.
 - ii. Guided Practice: Students begin drafting their writing using scaffolding from Gemini AI materials under the guidance of a lesson plan.
 - iii. Independent Practice: Students complete their writing independently while continuing to refer to the material structure provided by AI to maintain consistency in the quality of their writing.
 - b. Control Class (Conventional Learning): Students in the control group receive the same contents and topics but use conventional learning material without the help of Gemini AI. The writing process is done manually using textbooks or teacher instructions as usual.
3. Post-Experiment Stage (Post-Test)

After the series of learning sessions were completed, both groups were given a post-test (final test). The instrument used had the same level of difficulty and format as the pre-test. After that, the results of this post-test served as the primary quantitative data to see if there was a significant

advancement in writing proficiency occurred within the experimental group in comparison to the control group.

4. Documentation and Score Collection

During and after the research, the researcher systematically collects all supporting data. Student worksheets are collected for analysis. Then, scores from pre-tests and post-tests are entered into data tables based on analytical assessment rubrics (Appendix 9). Documentation in the form of lesson plans and examples of Gemini AI synthesis outputs is archived to ensure transparency of research procedures.

5. Data Finalization

All scores collected from both classes are organized chronologically. This numerical data is then prepared for processing through statistical analysis to evaluate the research hypothesis, specifically to ascertain whether the utilization of Gemini AI structured materials exerts a significant influence on learners' writing.

E. Data Analysis Technique

The data gathered from the pre-test and post-test were evaluated utilizing statistical techniques to test the research hypothesis. As noted by Creswell (2012), quantitative data analysis in experimental research focuses on comparing group scores to show whether the treatment has a statistically significant impact. The data analysis procedures in this study are as follows:

1. Descriptive Statistics

The first step is to describe the raw data to offer an overview of student achievement. The descriptive statistics used include the following calculations:

- a. Mean: To see the average writing achievement of students in both classes.
- b. Standard Deviation: To see the distribution or variation of student scores within the group.
- c. Maximum and Minimum Scores: To see the range of student achievement in the writing test.

2. N-Gain Analysis

N-Gain analysis was conducted to ascertain the level of students' score advancement between the pre-test and post-test in both groups. The resulting N-Gain values were utilized to evaluate the level of efficacy of the intervention based on the progression of student scores following the instructional period. The N-Gain categories were classified into high, medium, and low categories.

3. Prerequisite Test

Before conducting hypothesis testing, the data had to satisfy several statistical assumptions so that the appropriate statistical analysis could be determined (Sugiyono, 2017).

- a. Normality Test: This assessment was deployed to verify whether the dataset adhered to a normal distribution. The investigator utilized the

Shapiro-Wilk test via SPSS software, given that the sample size within each cohort comprised fewer than 50 participants.

- b. Homogeneity Test: This test was implemented to ascertain whether the variances between the experimental and control cohorts displayed homogeneity.

4. Within-Group Analysis

Within-group analysis was conducted to evaluate whether a meaningful disparity existed between the pre-test and post-test scores within each respective group following the instructional period. If the data demonstrated normal distribution, the Paired Sample T-Test was utilized. Conversely, if the data were not normally distributed, the Wilcoxon Signed Rank Test was applied.

5. Hypothesis Testing

To evaluate whether a meaningful disparity existed between the experimental group and the control group, a comparison test was performed. Creswell (2012) asserts that a mean comparison test is appropriate for contrasting the outcomes of two groups within experimental investigations.

The Independent Sample T-Test was utilized if the data fulfilled the normality assumption. On the other hand, if the data were not normally distributed, the Mann-Whitney U Test was applied as a non-parametric alternative.

6. Decision Making

Decisions regarding the hypotheses were made based on the significance value (p-value) with the following criteria (Sugiyono, 2017):

- a. If the significance value (Sig. 2-tailed) < 0.05 , the null hypothesis (H_0) was rejected and the alternative hypothesis (H_a) was accepted. This indicates that there was a significant effect of using Gemini AI-synthesized structured materials on students' writing skills.
- b. If the significance value (Sig. 2-tailed) > 0.05 , the null hypothesis (H_0) was accepted, indicating that there was no significant effect.

To prevent manual calculation errors and guarantee the precision of the outcomes, all statistical analyses in this investigation were performed using SPSS software (Appendix 10).

D. Validity and Reliability of the Research

To ensure the validity of the research instruments, expert validation was conducted before the treatment. The pre-test and post-test tasks, scoring rubric, lesson plans, and Gemini AI-synthesized structured materials were reviewed by English teachers at SMA Negeri 1 Ngimbang (Appendix 12). The purpose of this validation was to ensure that the instruments were appropriate for eleventh-grade students, aligned with the curriculum, and suitable for measuring the objectives of the study. Based on the validation results, the instruments were considered appropriate for use in the research.

To maintain the reliability of the research data, a standardized analytical scoring rubric adapted from Weigle (2002) was utilized. The students' writing performances in both the pre-test and post-test were assessed

by two raters: the researcher and an English teacher at SMA Negeri 1 Ngimbang. The rubric evaluated five components of writing, namely content, organization, vocabulary, language use, and mechanics. The final score for each student was obtained by calculating the average of the two raters' scores (Appendix 9). The involvement of two raters and the use of the same scoring rubric were intended to enhance scoring consistency and minimize subjectivity in the assessment process.