CHAPTER III
RESEARCH METHOD

3.1. Research Design

The researcher utulized experimental design. There were two cohorts implemented such as experimental cohort and control cohort. Pre-test was done before the experimental and control cohort received the treatment, and post-test was done after the treatment period. 8 meetings were held to complete this research. The use of a WhatsApp group to implement MCL with MSS . The experimental study model is shown below:

Table 1. The model of experimental study

Group	Pre-Test	Treatment	Post-Test	Questionnaire
Experimental	Writing Test	MCL (Group Magnet Summary)	Writing Test	Closed-ended questionnaire
Control		Individual Traditional Method		

3.2. Research Setting

This research was administered at MA Al-Balagh Lamongan . At this school, there were six classes. The researcher chose this institution for the following reasons: (1) There has never been a researcher at this school who has conducted a study with the same title. (2) the researcher's boarding house is conveniently located near the school. Meanwhile, the research subject is the tenth grade of MA Al - Balagh Lamongan.

3.3. Reseach Population, Sample and Technique of Sampling

The population consisted of 50 students in 10 grade at MA Al-Balagh Lamongan, there were 25 of whom were enrolled in the natural science program and 25 in the social science program. The experimental group (N 25) with 11 male and 14 female students, while in the control group (N 25) with 13 male and

12 female students. The researcher determined the experimental class and the control class from the existing population.

Following the selection of the sampling, the experimental cohort (X-IPA) and control cohort (X-IPS) were chosen. The sample for this study was selected using the total sampling methodology. As a result, they are all used as research participants or respondents. According to sugiyono (2018) states that total sampling is a sampling method which utilises all members of a population as samples. This is commonly done when the population is limited (less than 30 respondents) or when the study intends to make broad generalizations with very small margins of error (p.125).

3.4. Research Instruments

The design of this study conducted a quasi-experimental design. As a result, he would gain data collection tools using a writing test and a questionnaire.

3.4.1. Tests

A test is a series of accomplishments and additional resources used to evaluate one's or a group's capabilities, abilities, information, and competences. (Suharsimi Arikunto, 2006). A writing assessment will be given to both the experimental and control groups.The student's textbooks are used to create the writing test. Essay is the writing test format because it is easy to administer, quick to score, and quite familiar to students.

The test are created by the researcher based on a syllabus that emphasizes on writing ability. The descriptive text writing score will be used by the researcher to assess the students' writing abilities. The study employs a rubric assessment adapted by Brown (2007), which classified the five characteristics of writing production into five categories: (1) content; (2) grammar; (3) organization; (4) vocabulary; and (5) mechanics.

3.4.2. Questionnaire

The researche conducted the closed-ended survey. The experimental group received a questionnaire following the course of treatment.

Larry Cristensen (2004) as cited in (sugiyono, 2018) states:

> *"a questionnaire is self-report data collection instrument that each research participant fill out as part of a research study. Researcher use questionnaires so that the can obstain information about the thoughts, feeling, attitudes, belieft, values, perceptions, personality, and behavioral intentions of research participant. In other words, researchers attempt to measure many different kinds of characteristic using questionnaire"*. (p.193).

The questionnaire's objective is to collect data about students' perceptions' toward participating in Mobile Collaborative activities using MSS to enhance students' writing performance. The questionnaire consists of the 20 items that make up the 5 points Respond Scale in the questionnaire: Strongly disagree on 1; disagree on 2; Agree on 3; Strongly agree on 1.

3.5. Procedure of the Research

The course would be held 8 meetings, each of which lasted for the same length of time as an English class. English is taught for 2 hours a week or 4x45 minutes, with a schedule of once a week (2x45 minutes for each meeting) to collect data from both tests and questionnaires.

3.5.1 Pretest

First meeting, the researcher asked the students to write descriptive text. They write text based on a descriptive text structure of 2 paragraphs or more with a minimum of 100 words. They are free to silect one of the themes in the pretest questions. With a time allocation of 90 minutes. All submission must be original, not reprinted from nternet or othersource.

3.5.2 Treatment

After the students did the pretest, the researcher started using mobile learning via Whatsapp in teaching the class how to write descriptive texts as well

as explaining the magnetic summary strategy to students. Treatment is given in 8 meetings. In the second meeting , he sent descriptive text to the group by explaining the descriptive text definition and giving examples, followed by a magnet summary strategy . The writing aspect include content, organization, grammar, vocabulary and mechanic. This assessment criterion is adapted from Brown (2007). Towards the end of the meeting, students were given exercises to train them to identify good and bad writing by working in a group. At final meeting, the researcher explained the plans for the next meeting, where students were introduced to Whatsapp group activities. He supervised every student activity by online. Likewise for the fourth meeting until the seventh meeting.

Experimental Group (EG) :

a) Teacher created WhatsApp group consisting of 25 learners in group.
b) Class were divided into 5 groups.
c) Learners were enganged in process of teaching learning individually
d) The learners discussed questions asked by the teacher on the topic in the group via WhatsApps group or video conference.
e) Every week, The teachers and learner discussed the answers of the questions via WhatApps group.
f) The teacher gave feedback and comment.

Control Group (GC) :

a) Each learner was assigned to individual (face to face)
b) No grouping needed
c) No treatment in MCL and MSS but using teacher's individual strategy
d) The teacher explained the subject matter.
e) The teacher discussed students' answer.

Table 2. The Following Research Treatment Conducted By Researcher

No	Date	Activities
1.	November 15, 2022	Pre test X-IPA (Experimental Class)
2.	November 16, 2022	Pre test X-IPS (Control Class)
3.	November 22- Desember 27, 2022	Treatment
4.	Desember 28, 2022	Post test X-IPA (Experimental Class)
5.	Desember 29, 2022	Post test X-IPS (Control Class)
6.	Desember 31, 2022	Distribution of questionnaires to students X-IPA (experiment Class)
7.	Desember– January, 2022	Data processing and analysis

3.5.3 Post-test

Learner are asked to compose descriptive text by selecting one of the themes in the post-test, which is not very different from the pre-test. But it must be different from the essay during the pre-test. They wrote a descriptive text with 2 paragraphs or more with a minimum of 100 words. Time to work on 90 minutes. Additionally, the writing cannot be plagiarized from the internet or other sources.

3.6. Techniques of Data Collection

The instruments data which utulised are writing test and questionnaire. MCL with MSS implemented in experimental group. The Traditional Method (Individual) is administered to the control group. Additionally, students in the experimental class are required to complete a survey about their opinions of the implementation of MCL with MSS via WhatApps group for English learning.

3.7. Technique of Data analysis

To analyze the data, the researcher utulised IBM SPSS (Statistic packag for Social Science) Statistic version 24 to analyze the data for determining the difference between unpaired group in posttest control and experiment. Yet, paired sample t-test was utulised calculated in same group (pre and posttest experiment to know wether there was improvement after conducting treatment. The information

was then displayed utilizing the procedures and methods described below:

3.7.1. Scoring of descriptive Writing Test

The researcher welcomed English teachers from SMK Negeri 1 Sambeng and MAN 2 Lamongan, both of whom had acquired master's degrees in English education and had more than three years of teaching experience, to grade the descriptive writing. Analytical writing rubric was used to evaluate descriptive writing performance.

When evaluating a writing test, there are several levels of scoring such as :, , excellent, good, fair, poor and very poor. Students can do better than them score between 81 and 100. The students got scores range from 66 to 80 are well-classified. The classification is considered fair if the student obtains between 56 to 65. Misclassification of students who get grades 47-55. If students with scores below 46 are misclassified. The following categories of descriptive writing as follows:

Table 3. The Categories of Scoring the descriptive writing

Score interval	Classification
81 – 100	Excellent
66 – 80	Good
56 – 65	Fair
47 – 55	Poor
Below 46	Very Poor

(Source : MA Al – Balagh Bulutigo. In the Academic Year of 2022-2023)

3.7.2. Data Description

Two analysis was conducted on the data description

1) Frequency Distribution

The distributions of frequency data gain the students' scores and frequency. The students' pretest and post scores are used to generate the frequency

data distributions of control and experiment. A table analysis is then used to display the frequency distribution of the data.

2) Descriptive Statistics

The descriptive statistic is gained the minimum, maximum, standard deviation, and standard error of the mean score. The descriptive statistics are generated from the students' scores in pretest or pretest both control and experimental cohort.

3.7.3. Pre-requisite Analysis

3.7.3.1. Test of Normality

To determine whether the collected data had a normal distribution, testing was carried out. The Shapiro Wilk test was used. The data gained from students' pre-posttest both experiment or control group. The data can be indicated not normally distributed, if the Sig. < Alpha Research (0.05). Otherwise, the data can be normally dispersed, if the sig > Alpha Research (0.05) (MJurnal, n.d.)

3.7.3.2. Test of Homogenity

The data gained from students' pre- and post-test scores to determine the homogeneity score. If significance value (probability) on the Levene's Statistic is higher than 0.05, the variables variance is homogeneous, in both the experimental and control groups.

3.7.3.3. Test of Hypothesis

T-test was utulised to compare the students' mean score in unrelatedgroup both control and experimental group The following is the research's hypothesis:

1. If the T-test score is greater than the t-table. It indicates that there is a significant effect of MCL with MSS implementation on students' writing performance via WhatsApp. So, H1 is accepted
2. If the T-test score is less than the T-table score. This indicates that there is no significant effect of MCL with MSS implementation on students' writing performance via WhatsApps . So, H0 is rejected.

3.7.4. Questionnaire Responds analysis

The second purpose of this research was to explain EFL learners' perceptions toward MCL with MSS. The percentage of students who responded to the questionnaire calculated after it was conducted the treatment to the students in the experimental group. There were categorization of chosen scale: Strongly agree, Agree, Disagree, and Strongly Disagree.

3.8. Test Validity and Reliability

3.8.1. Validity

Testing can be said a good must match its validity requirements. The experts were then given the items to check that the test's content validity. By checking off items on a checklist for validating the English-writing test, the experts were asked to validate and evaluate the test.

To gauge the test's content validity, two experts were administered the first test revision to assess each item's appropriateness in evaluating writing abilities (Fattah, 2015). They were also asked to assess the test as a whole in terms of 5 items are assessed, including: *Instructions, topics, timing, content and content appropirate* . In summary, writing tests are valid and suitable for data collection take an exam content validity and construct validity were tested.

3.81.1. Construct Validity

The validity was used to determine whether or not the test is in line with the theories that support the information presented. The researcher utilized a grading rubric developed from Brown (2007) to grade students' descriptive writing tests. Brown (2004) defines that legitimizing large-scale standardized tests of skillfulness is complicated by construct validity. In addition, the researcher asked the raters to measure the instrument test format. They were English teacher English Teachers from SMK Negeri 1 Sambeng Lamongan and MAN 2 Lamongan who had pursued master degree of English education. The five item are checked such

as : instruction, time allotment, topic, content , and rubric.

3.8.1.1. Content Validity

In order to guarantee that the test content correctly shows the class or fields of the positions or fields existing, content validity is implented. This is accomplished through professional judgments regarding the test's relevance and sampling to a particular domain. Coverage and representativeness are more important than response patterns or scores. (Cohen, et al. 2007). The researcher then asked two raters to evaluate the instrument test format. According to Fraenkel, et al (2012) stated that the content and format of the instrument are referred to as content validity. In this case, a book for Senior High School was utilised. The writing test specification is displayed in table 3.

Table 4. Test Specification of Writing Test

Basic Competence	Learning Objective	Indicator Competence	Material	Test types	Item Number
3.1. Distinguishing social functions, text structures, and linguistic elements of several oral and written descriptive texts by giving and asking for information related to tourist attractions and famous historical buildings, short and simple, according to the context of their use. 4.4. Descriptive text 4.4.1. Capturing contextual meaning related	After discussing the material the learners are hoped to be able to : 1. To Identify descriptive generic structure 2. To identify descriptive language feature 3. To generate any ideas by using magnet word. 4. To Make a summary composition by using magnet words to generate MSS correctly. 5. To compose descriptive text by composing the simple short essay written text	1. Learners can correctly generate a summary correctly from magnet words generated by MSS activity. 2. Learners can compose short essay written text of descriptive text 3. Learners can Identify the descriptive generic structure 4. Learners can comprehend descriptive language text.	Descriptive text	Written test	1

to social functions, text structure, and linguistic elements of descriptive, spoken and written, short and simple texts related to tourist attractions and famous historical buildings. 4.4.2. Compose oral and written descriptive texts, short and simple, related to tourist attractions and famous historical buildings, taking into account social functions, text structures, and linguistic elements, correctly and in context		5. Learners can generate ideas using magnet word.			

3.8.2. Reliability Test

After validating the instrument validity, the next stage is to evaluate the reability. The grade at which point a test yields compatible results or scores is popular as reliability. Fraenkel, et al. (2012) says that reliability is the compatibility of the scores acquired, or in what way or manner correct they are each human from individual mechanism presidency to the next and from one set of parts to the next.

Raters (expert judgements) are promoted to evaluate if the rubric and instrument were reliable or not. To determine wether or not the test was reliable, two expert judgement were used. English teachers from SMK Negeri Sambeng Lamongan and MAN 2 Lamongan provided the expert opinions. According to

Brown (2004) Inter-rater reliability is a problem that frequently affects classroom teachers. Because of unclear scoring criteria, weariness, bias against specific "good" and "poor" students, or just plain negligence. Inter-rater reliability can be known as the level of agreement among appraisers.

There are five aspects evaluated such as; instruction, time allotment, content, and content appropriateness (rubric). In nutshell, the data collection from writing test was reliable. Since its reability indicates on medium level.

The Guilford rule categorization (1956) are then used to interpret the reliability coefficient in the table as follows:

Table 5. Reliability Coefficient Interpretation

(Source: Guliford , 1956)

Reliability Coefficient	**Interpretation**
$0,80 < r_{11} \leq 1,00$	Reliability (Excellent)
$0,60 < r_{11} \leq 0,80$	Reliability (Good)
$0,40 < r_{11} \leq 0,60$	Reliability (Fair/medium)
$0,20 < r_{11} \leq 0,40$	Reliability (Less)
$0,00 < r_{11} \leq 0,20$	Reliability (Low)
$r_{11} \leq 0,00$	Unreliable

UNIVERSITAS MUHAMMADIYAH
GRESIK