CHAPTER III
RESEARCH METHODOLOGY

3.1 Research Method and Design

The method used was to evaluate the impact of a discovery learning teaching intervention through student questionnaires, tests completed before and after a specific teaching sessions. The research method was pre-experimental method, which is the assessing of posttest and pretest to analyze the change that occurs (Creswell, 2012).

The research design used in this research is the one-group pretest-posttest design which used to assessing the outcome of a group after a certain treatment. Treatment used in this research was the implementation of discovery learning assisted by Legends of Learning (LOL), a web educational games. The sampling techniques that used was convenience sampling, which defines as a group of individuals who are available for study (Fraenkel, 2011). The research design plan in this study shown in Table 3.1.

Table 3. 1
One-group pretest-posttest design

O	X	O
Pretest	Treatment	Posttest

(Fraenkel, 2011)

The treatment was applied only to one experimental group without a control group. The purpose of this research is to find out whether the treatment that applied could improve students' concept mastery and curiosity after implementing discovery learning assisted by LOL web-educational game.

3.2 Population and Sample

The research and data taken was gathered from a Junior High School in Cimahi, Indonesia. The placement was originally set by the campus and this school were chosen for the researcher to conduct the teaching practice as well as gathering the data. In this school, the science department divided into Biology and

Physics. In line to the targeted topic, which is "Light", the researcher take the Physics class in order to implement the treatment to the respondents. This research involved 50 respondents from three different classes in 8th grade that never received any lessons of the topic.

Table 3. 2

Detail of respondent number

Class	Total Students	Pre-test Respondents	Post-test Respondents	Both Pre- and Post-test Respondents
A	29	22	12	12
B	29	25	17	17
C	29	22	21	21

As can be seen in the table, the total amount of the original respondent that involved was 87 students from three different classes. The respondents are mostly 13 until 14 years old. Due to data completion, the only amount of students who complete both pretest and posttest data as 50 students from those three classes of 8th grade level. Therefore, the researcher only analyze the data from 50 respondents with 22 (44%) male students and 28 (66%) female students.

3.3 Operational Definition

In order to avoid misconception about this research, some operational definitions are needed in this research.

a. Discovery learning is a method where the learner draws on their own experience and prior knowledge. This learning method has a few stages which are: (1) Stimulation, (2) Problem Statement, (3) Data Collection, (4) Data Processing, (5) Verification, and (6) Generalization. This method involves instruction through which students interact with their environment by exploring and developing new knowledge (Castronova, 2002). On the other hand, LoL is the web- educational game which used to assist the implementation of Discovery Learning, specifically the 3rd and 4th stage (Data Collection and Data Processing).

b. According to Anderson and Krathwohl (2001), concept mastery is defined as the competence of students to reach expected cognitive level. In this research, the cognitive level that will be measured as an improvement of concept mastery only consisting of four levels, Remembering (C1), Understanding (C2), Applying (C3), and Analyzing (C4). The assessment was in a form of 20 multiple choice questions each for pretest and posttest.

c. In this research, students' curiosity was assessed using a set of statements adopted from a journal titled Science Curiosity in Learning Environments (SCILE) by Weible and Zimmerman (2016), they divided the statements into three aspects, which are science, stretch and embrace. This assessment instrument was used before and after the teaching intervention.

d. Light topic that being implemented in this research is one of the chapter from IGCSE Physic book which also used in the school whereas the data taken.

3.4 Assumption

The foundational assumptions of this study are:

a. Discovery Learning as instructional teaching method could help improve Indonesian student competency in science learning

b. Legends of Learning (LoL), a web-educational game can enhance students' ability in scientific understanding and curiosity.

3.5 Hypothesis

The hypothesis that is tested in this study:

a. H_0: There is no enhancement in students' concept mastery and curiosity in learning about visible light waves and its properties after the implementation of discovery learning supported by the LoL web-educational game.

b. H_1: There is an enhancement in students' concept mastery and curiosity in learning about visible light waves and its properties after the implementation of discovery learning supported by the LoL web-educational game.

3.6 Research Instrument

There are three types of instrument that are used in this research which are an objective test to measure student concept mastery, a questionnaire to measure student curiosity and an observation sheet.

3.6.1 Student Concept Mastery

The first test item is a set of 20 multiple choice questions. The questions were prepared with cognitive domain considerations and distributed based on each sub-topic. The question distributions are shown in Table 3.3.

Table 3. 3

Test item of students' concept mastery

No	Indicator	Test Item
1	Remembering	3, 5, 14, 16, 18
2	Understanding	1,2,4,6,8,12, 17
3	Applying	9, 15, 19, 20
4	Analyzing	7,10, 11, 13

The questions were judged by experts and tested for the validity of the items with a different group of students from the actual respondent groups. The group consisted by 20 students of 8th grade level from random schools that already learned the Light topic. The software ANATES was used in the process of knowing the validity of the instrument.

The ANATES has several aspect to analyze the data, which are validity, reliability, level of difficulty and discriminating power.

a. Validity

Validity refers to how correct, meaningful and valid is the specific conclusion that is collected (Frankel, 2011). The instrument that will be used must be valid in order to obtain a meaningful result from the study. The interpretation of validity is given in Table 3.4.

Table 3. 4

Interpretation of validity

Reliability Test Coefficient	Classification
0.00 - 0.20	Poor
0.20 - 0.40	Fair
0.40 - 0.60	Moderate
0.60 - 0.80	Good
0.80 - 1.00	Very Good

(Suharsimi, 2006)

b. Reliability

Reliability refers to the consistency of scores or answers from other instruments and from one item to another item, and from one set of items to another set. The classification results of reliability test is depended on the value of r, it is shown in Table 3.5

Table 3. 5

Classification of reliability test

Value of r	Interpretation
0.00 - 0.20	Very Low
0.20 - 0.40	Low
0.40 - 0.60	Enough
0.60 - 0.80	High
0.80 - 1.00	Very High

(Halim, 2010)

c. Difficulty Level

Difficulty is the mean score of test items, corresponded to the proportion of students who answer the item correctly. Arikunto (2006) stated that question which is not too easy nor too difficult could be defined as an good question. The value of difficulty index indicates the difficulty level. The interpretation of difficulty level is given Table 3.6.

Table 3. 6

Interpretation of difficulty level

Value of Difficulty Index	Interpretation
0.0 - 0.30	Difficult
0.30 - 0.70	Moderate
0.70 - 1.00	Easy

(Suharsimi, 2006)

d. Discriminating Power

Discriminating power analysis is the questions' ability to distinguish between the high and low achieving students. If the question is not good enough, then it could be answered by both low and high achieving students. In other case, if the questions could not be answered by both low and high achieving students, it also indicates that the question had no discrimination power. Suharsimi (2006) stated the interpretation of discriminating power is provided in Table 3.7.

Table 3. 7

Interpretation of discriminating power

D=	Quality	Recommendations
> 0.39	Excellent	Retain
0.30 - 0.39	Good	Possibilities for Improvement
0.20 - 0.29	Medium	Need to check/review
0.00 - 0.19	Poor	Discard or review in depth
< -0.01	Worst	Definitely discard

(Suharsimi, 2006)

The test item which was analyzed using ANATES is the test for assessing student concept mastery. The test item consist of 24 questions. The results of the tested test items are shown in Table 3.8, detailed in Appendix B.1.

Table 3. 8
Analysis of test item for student concept mastery

Question Number	Discriminating Power	Difficulty Level	Validity	Status
1	Excellent	Moderate	Low	Used
2	Medium	Easy	Very Low	Used
3	Medium	Moderate	Very Low	Used
4	Excellent	Moderate	Enough	Used
5	Excellent	Moderate	High	Used
6	Excellent	Moderate	High	Used
7	Poor	Easy	Very Low	Revised
8	Medium	Easy	Low	Used
9	Excellent	Difficult	Enough	Used
10	Medium	Easy	Very Low	Used
11	Excellent	Moderate	Enough	Used
12	Medium	Easy	Very Low	Used
13	Excellent	Moderate	Enough	Used
14	Medium	Easy	Enough	Used
15	Medium	Easy	Low	Used
16	Excellent	Easy	Low	Used
17	Poor	Moderate	Very Low	Rejected
18	Poor	Difficult	Very Low	Rejected
19	Medium	Easy	Low	Used
20	Medium	Easy	Low	Used
21	Worst	Moderate	Very Low	Rejected
22	Medium	Moderate	Very Low	Rejected
23	Excellent	Difficult	Low	Used
24	Poor	Moderate	Very Low	Revised

The result from ANATES software analysis toward the concept mastery test item resulting 4 questions rejected, which are question number 17, 18, 21 and 22.

Two questions were revised, which are question number 7 and 24. There was also expert judgement considerations, and the researcher finally decided 20 questions as final objective test item set.

3.6.2 Student Curiosity

The curiosity improvement rubric is in a statement set. A five-Likert scale model is used to determine each statement with five scale level; Always (5), Often (4), Sometimes (3), Not Often(2), and Never (1). This instrument was used to assess the student before and after the treatment. It was consisted of three aspects that has been developed and adapted by Weible and Zimmerman (2016), the statement distribution for each aspects given in Table 3.9.

Table 3. 9

Questionnaire statements

Aspect	Statement
Science	1, 2, 3, 4, 5, 6, 7, 8
Stretch	9, 10
Embrace	11, 12

The 12 questions of SCILE was chosen based on the strong response and correlation seen in our big dataset with responses from over 650 young people, with the scale's reliability result on Cronbach's α of 0.91 (Weible, 2016).

3.7 Research Procedure

In order to organize this research with a proper plan, there are several steps that should be followed: Preparation, Implementation, and Completion.

a. Preparation Stage

The first stage is about planning and preparing everything, as detailed below:

1) Formulation of the research problem
2) Decide the variables as the focus of the research
3) Literature review to analyze curriculum objectives, current scientific problems, discovery learning, Legends of Learning, concept mastery, curiosity, and of course the chosen topic, 8th grade level Light topic from

Physics subject.

4) The arrangement and preparation for proposal seminar
5) Revision of the proposal based on comments and suggestions by the lecturer and examiners
6) The preparation and development of instruments, including an objective test for measuring concept mastery, a questionnaire statements to measure curiosity, teaching materials for the Light topic, lesson plans, and worksheet to support the treatment implementation.
7) Instrument revision after receiving feedback from experts.
8) Conduct additional test to the students which already learned about the Light topic before, so the result data can help the researcher to analyze the instruments' validity.
9) Analysis using ANATES software to identify the validity level
10) The revision and consideration of the concept mastery objective test instrument that will be used and rejected.

b. Implementation Stage

In this stage, the valid instruments is used as follows:

1) Pretest
2) The implementation of the treatment to the population
3) Posttest conducted in order to measure the enhancement level of concept mastery and curiosity

c. Completion Stage

In this last stage, the collected data is analyzed by the researcher, including these steps:

1) Analysis and calculation of the data
2) Connect the result to the literature reviews and prove the points
3) Answer the hypothesis
4) Conclude the result and findings
5) The final arrangement of the whole result as thesis

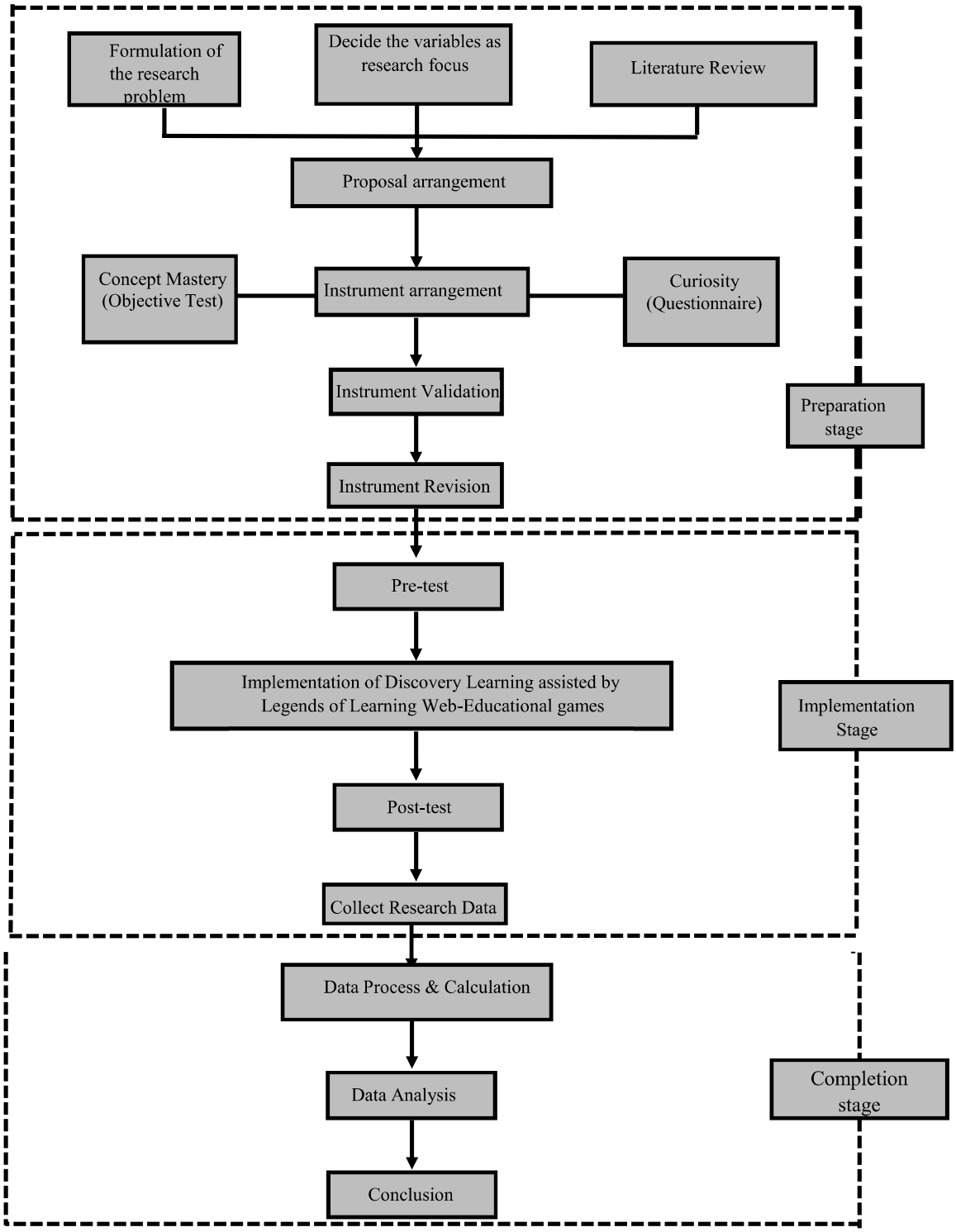

Figure 3. 1 Research Procedure Flowchart

3.8 Data Analysis

All of the data was analyzed both qualitatively and quantitatively. The explanation of the data processing techniques as follows:

3.8.1 Student Concept Mastery

The analysis of objective test to measure student concept mastery was done by Microsoft Excel and SPSS for windows version 25.0. The process of the data calculation is explained as follows:

a. Scoring

The result of pre-test and post-test test items were calculated to know the scores from both tests. From 20 selected multiple choice questions, students' score were all converted into percentage score. All of the data saved in Microsoft Excel.

b. Normality Test

This test is to check the normal distribution of a set. According to Frey (2018), a normal distribution is a hypothetical symmetrical distribution used to make comparisons among scores and to make other kinds of statistical decisions. The shape of this distribution is often referred to as "bell shaped" which implies the majority of scores lie close to the center of the distribution, and the data is normally distributed. To test the normal distribution, SPSS statistic version 25.0 was used.

After the researcher find out the normality of the data distribution, the next test would be deciding whether to use paired t-test, or wilcoxon-signed rank test. Both of the test have similar function, which is to prove whether the hypothesis is accepted or not. If the data is normally distributed, then it will be continue to be tested using paired t-test, and if the data is not normally distributed, then it will continue to be tested using wilcoxon-signed rank test.

c. Homogeneity

This test aims to determine the same variance of the both samples. If the data has the same variance, means the populations identified as homogeneous variance,

and if it wasn't, the populations determined as the heterogeneous variance.

d. Paired T-Test

This t-test was used to test the significant difference between two groups which has normal distribution and analyze the effect of the independent variable to the dependent variable. It will prove which hypothesis is accepted and which one is rejected as the result of the research.

e. Wilcoxon Signed-Rank Test

The Wilcoxon signed-rank test is a non - parametric test which also used to test which hypothesis is accepted and rejected. This is an alternative when the distribution of the difference between two samples' means assumed to not distribute normally.

f. Gain and Normalized Gain

The normalized gain (N-Gain) is a rough measure of the effectiveness of a course in promoting conceptual understanding (Hake, 1998). This test is often used to measure the improvement or the gain from pretest to posttest data using the defined formula below:

$$< g >= \frac{\%S_f - \%S_i}{100 - \%S_i}$$

Description:

<g> = Normalized gain
Sf = Posttest score
Si = Pretest score

The enhancement of the data categorized based on the interpretation of N-gain at Table 3.10.

Table 3. 10

Interpretation of N-Gain

N- gain Score	Category
(<g>) > 0.7	High
0.7> (<g>) > 0.3	Medium
(<g>) < 0.3	Low

(Hake, 1998)

3.8.2 Student Curiosity

Student curiosity data is tested using a questionnaire using five-Likert scale is analyzed first using Microsoft Excel to calculate the score of each students. After that, the data was calculated to know the pretest and posttest results. In order to know the enhancement, the N-Gain score was calculated. The average score of each students to know the difference in the number of students in each aspect before and after the learning process is also calculated. After that, the enhancement being categorized based on the interpretation of N-Gain score shown in table 3.10.

Beside on analyzing the whole enhancement, the 12 items in the questionnaire are also analyzing the scientific curiosity through three correlated aspects, which are stretching, embracing and science practices. The N-Gain for each aspect were also calculated to see the enhancement and being categorized based on the interpretation of N-Gain score.