CHAPTER III
RESEARCH METHODOLOGY

3.1 Research Design and Methodology

This research used a quantitative technique. According to Creswell (2012), quantitative research is any kind of study that uses numerical data collection and mathematically based methodologies (namely, statistics) for analysis in order to explain phenomena. Quantitative research has the advantage of being more organized, systematic, and transparent from start to finish, and it is also less influenced by field conditions. A variety of research designs, including survey, correlation, and experimental designs, can be used with quantitative techniques in a variety of investigations. Therefore, a survey design was used for this research. Researchers give a survey to a sample in order to describe the attitudes, opinions, behaviors, or features of the population. In the field of education, a survey design is a test that is administered without any particular treatment of the sample, data manipulation, or research settings. It focuses on data trends rather than providing thorough explanations or cause-and-effect correlations (Creswell, 2012).

The type of survey design employed in this research is cross-sectional research design. Cross sectional survey design, is a research design that collects data at one time on a sample (Creswell, 2012). Cross-sectional studies are observational studies that analyse data from a particular population at a single point in time. Cross-sectional surveys are a common option for the first stage of research because of their capabilities in identifying trends and gathering data effectively, despite their limits in proving cause and effect. Researchers can obtain a useful overview of the state of education now, pinpoint possible problem areas, and test more focused hypotheses for research by employing cross-sectional surveys (Kesmodel, 2018). In this study, the cross-sectional research methodology was used to collect data from eighth grade students at a certain point in time to gauge their grasp of themes related to cell topics. This method offers a quick glance at students' knowledge and misconceptions, offering insightful information for intervention and educational improvement plans.

In order to analyse this research, descriptive methods were used. When research is conducted with the goal of investigating and elucidating a phenomenon

through variable descriptions, it does not test pre-existing hypotheses, but instead presents the research findings as they are (Zellatifanny & Mudjiyanto, 2018).

3.2 Population and Sample

A sample is a subset of the population, whereas a population is the entire group of individuals with particular traits. Generally speaking, the majority of people consider a population's defining feature to be its geographic location (Thacker, 2020). The location of this research is a public junior high school in Bekasi, West Java. The school used the National Merdeka curriculum (Maryana, 2021). Data was obtained in the first semester of the 2023/2024 academic year with a specific focus on 8^{th} grade students who have been taught about cell topics.

Population of this research is 8^{th} grade students in 14 junior high school Bekasi. First, the sample of the validating test is 34 students, then the sample for the real test is 232 students who have studied cell material. Based on the rules and language guidelines of the participating schools, this study is being done in Indonesian. The language used is chosen to guarantee adherence to school policies and enable efficient communication with the students throughout the data gathering process.

Convenience sampling is the method used in this study. Convenience sampling is when a researcher chooses participants based on their availability to participate in the study (Creswell, 2012). Convenience sampling is a sampling technique wherein individuals are selected based only on their suitability as "convenient" sources of data for researchers (Lavrakas, 2008). Convenience sampling has several advantages. First, compared to other sampling methods, the researchers can choose the participants with less effort. Second, convenience sampling necessitates inexpensive participant selection on the part of the researcher. Third, because the sample is easily available, the researchers spend less time. Lastly, they are not required to compile a list of every member of the population (Golzar et al, 2022).

Useful data can be obtained if researchers carefully manage biases and uncertainty while employing a convenience sampling approach. By assessing and managing the sample's representativeness, including diversity, and utilizing additional data, convenience sampling biases can be reduced. First, by assessing

and managing the representativeness of the sample, the researchers can increase the research's usefulness and lessen biases in participant selection. Second, in order to gather a suitable cross-section of the target population, the researchers can further diversify convenience samples by distributing questionnaires at various times and locations. Lastly, another strategy to reduce bias and uncertainty is to use more data. By employing larger samples, the researchers could incorporate more data (Skowronek & Duerr, 2009).

3.3 Research Instrument

The instrument of this research uses an objective test, which measures students' misconceptions about cells, and is based on a four-tier diagnostic test. A four-tier diagnostic test is a development and modification of a three-level diagnostic test that consists of two sections: questions that assess students' cognitive comprehension and another section that explains the rationale behind the answers. The four-choice format of the questions is intended to limit the number of times students can provide a thoughtless response that is not according to the established context. Questions were first created through a preliminary study and subsequently transformed into a four-level diagnostic test to identify students' misconceptions. The tests consist of 20 multiple choice questions, which underwent a thorough validation phase. Thirteen questions were selected after validation to make up the final set of questions for the real test instrument. An overview of the subtopics addressed by the chosen questions is given in Table 3.1.

Table 3.1 Concepts of question distribution

Sub-Topic	Question
Cell & Microscope	1, 2, 5, 10, 12
Animal Cell & Plant Cell	3, 6, 9
Cell Specialization	4, 7, 8, 11, 13

The concepts presented to the students have been taught to them as part of their learning curriculum. The four-tier diagnostic test questions are crafted to gauge how well students grasp these concepts. The researcher can accurately

determine the students' misconception by including trap solutions, which are answers that seem plausible to students but are actually a diversion from the right answer. It is expected that students' responses would differ based on their comprehension level and possible misconceptions. The varying answers enable researchers to examine the frequency of particular misunderstandings and comprehend the rationale behind students' incorrect answers. One example question from the four-tier instrument is presented in Table 3.2.

Table 3.2 A Sample Question

No	Tier	Question
4	1st	Based on the number of constituent cells, bacteria are examples of living things?  source: alodokter.com A. Multicellular B. Unicellular C. Eukaryotic D. Hominids
	2nd	Are you sure about your answer? A. Sure B. Not Sure

No	Tier	Question
	3rd	What is your reason to choose that answer? A. Bacteria are organisms that have a nuclear membrane B. Bacteria are organisms that have many cells C. Bacteria are organisms that have a flat shape D. Bacteria are organisms that have one cell
	4th	Are you sure about your answer? C. Sure D. Not Sure
Scientific Concept		Based on the number of constituent cells, living things are divided into two, namely unicellular and multicellular. Unicellular is an organism that only has one cell and only needs one cell to carry out all its life activities. Bacteria are one example of unicellular organisms.

Students are given the question in Table 3.2 to determine their knowledge of cell specialization. In the first-tier, students are asked to identify what organisms bacteria belong to based on the number of constituent cells. Bacteria are an example of unicellular living things because they only have one cell and only need one cell to carry out all their life activities. Then, in the third-tier students are asked to choose the best reasoning for their answer in the first-tier question. In the second and fourth tier, students are asked to select whether they are confident of their answer.

3.3.1 Preliminary Study

In order to develop a four-tier diagnostic test, the researcher will create an initial twenty questions. These questions are intended to evaluate students' comprehension of cell subjects and pinpoint any possible misunderstandings. After the initial question set is created, the instrument will undergo two validation rounds to ensure validity.

The first step in doing research and creating a research model is a preliminary study. Problems can be analysed using preliminary study as an example for more in-depth investigation. Open-ended, two-level questions were employed in this topic's preliminary study. The first level consists of questions designed to gauge students' comprehension of the subject, while the second level consists of the rationale behind the response. Table 3.3 lists the open-ended questions that were used in the preliminary investigation.

Table 3.3 Preliminary Study Questions

Item	Part	Question
1	**1**	Are all living things composed of cells?
	2	Give a brief explanation of your previous answer!
2	**1**	What is the researcher's goal in making a microscope?
	2	Give a brief explanation of your previous answer!
3	**1**	Explain the difference between a light microscope and an electron microscope!
	2	Give a brief explanation of your previous answer!
4	**1**	Explain the difference between a monocular microscope and a binocular microscope! 
	2	Give a brief explanation of your previous answer!

Item	Part	Question
5	**1**	Based on the picture above, mention the function of the diaphragm on the microscope!

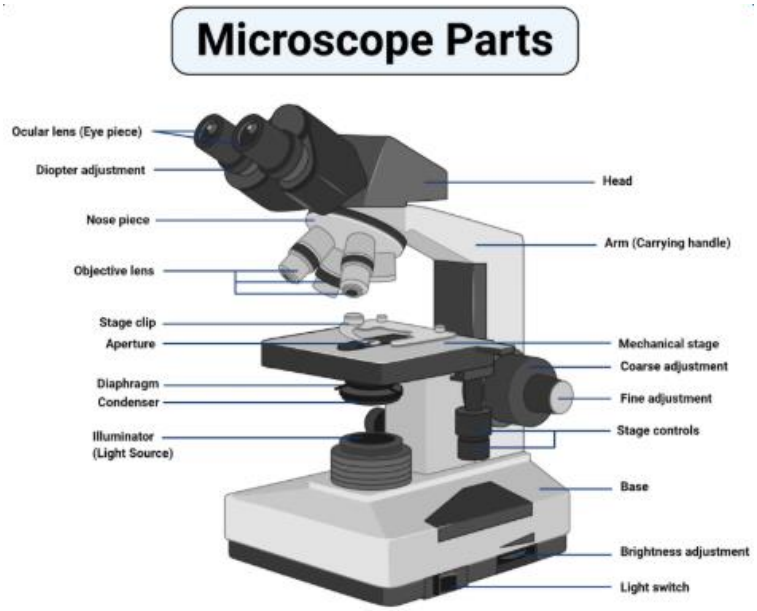


Figure: Parts of a microscope, Image Copyright © Sagar Aryal, www.microbenotes.com

	2	Give a brief explanation of your previous answer!
6	**1**	Based on the picture above, mention the function of the objective lens on the microscope!

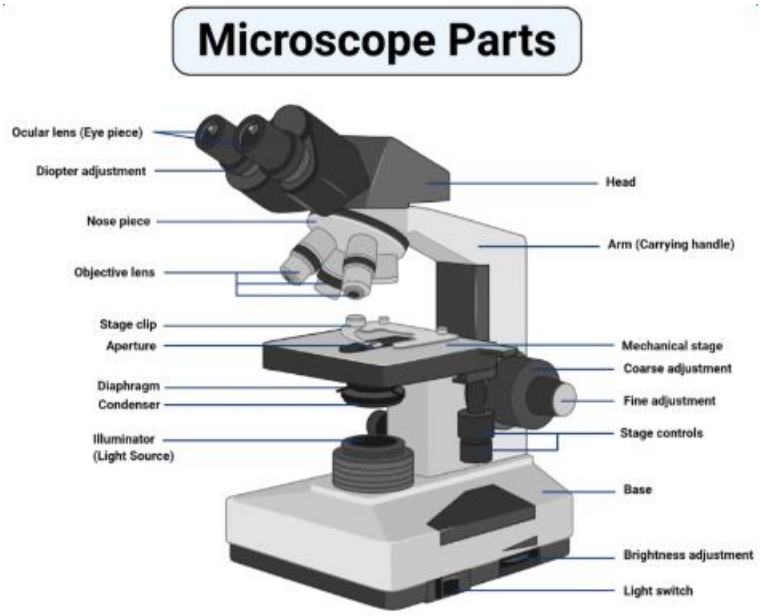
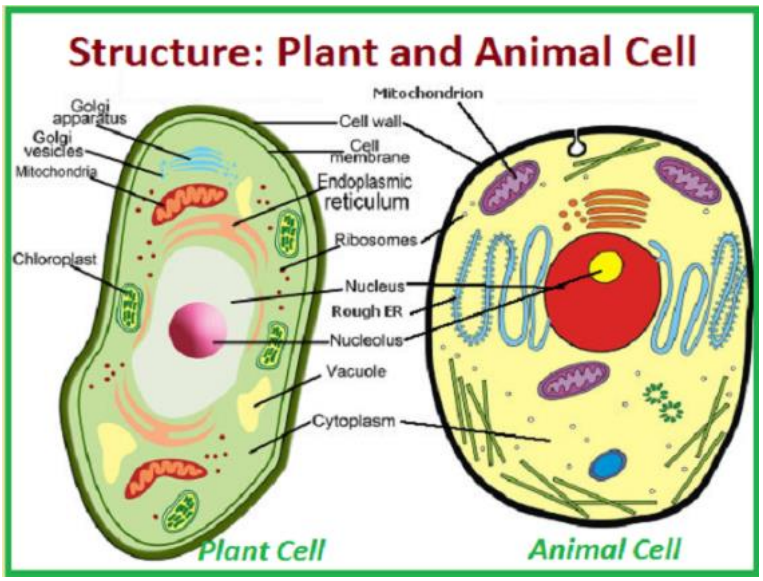
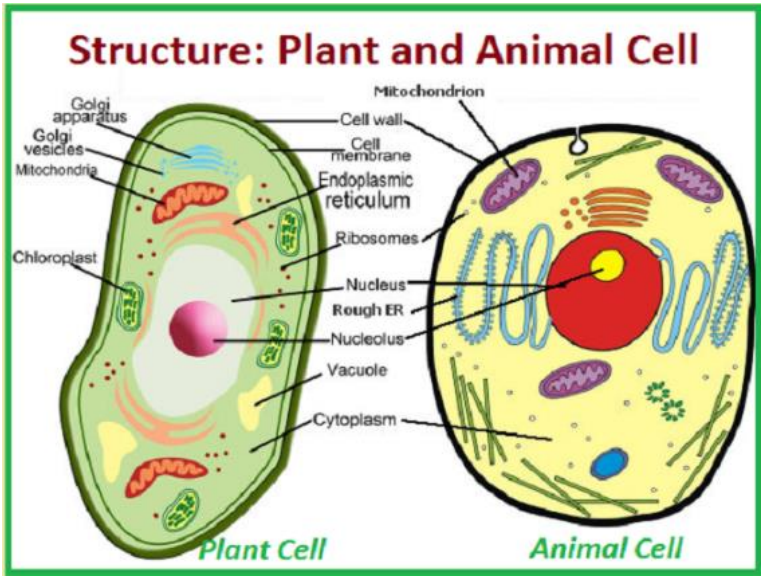
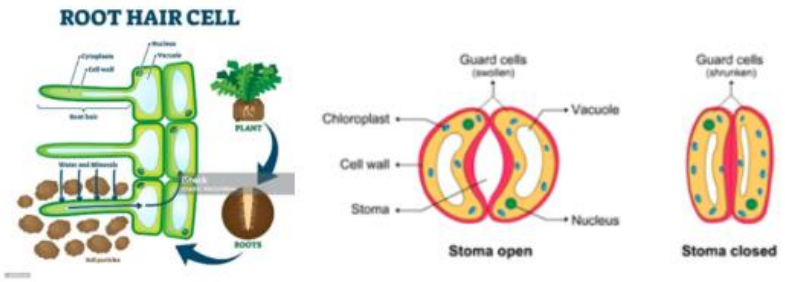


Figure: Parts of a microscope, Image Copyright © Sagar Aryal, www.microbenotes.com

	2	Give a brief explanation of your previous answer!

Item	Part	Question
7	1	Based on the picture above, mention the function of the nucleus!
		
	2	Give a brief explanation of your previous answer!
8	1	Based on the picture above, mention the function of the mitochondria!
		
	2	Give a brief explanation of your previous answer!
9	1	Explain the difference between plant cells and animal cells!
	2	Give a brief explanation of your previous answer!
10	1	Bacteria are examples of unicellular or multicellular living things?
	2	Give a brief explanation of your previous answer!

Item	Part	Question
11	**1**	Do the vacuoles in root hair cells and stomata work the same way?



a. Sel akar rambut **b. Stomata**

	2	Give a brief explanation of your previous answer!
12	**1**	Sort the stages into the correct body formers!

a. Sistem **b. Jaringan** **C. Organisme** **D. Sel** **E. Organ**

	2	Give a brief explanation of your previous answer!

Thirty-seven eight-grade students who had previously studied the cell topics participated in this preliminary study phase. Multiple-choice questions for the four-tier diagnostic test were created using the different answers that were gathered from the 12 questions, which were then grouped and generalized. Arbitrary misconceptions were present in some of the responses. Therefore, alternatives from various sources were employed in the construction of the diagnostic test questions. The amount of time spent on the test, students who have forgotten the material, and students who did not take the test seriously are some of the factors that led to less than ideal outcomes. Appendix A.1 contains general questions and answers from the preliminary phase.

3.3.2 Validation Test

The results were validated after being evolved into a four-tier diagnostic test. The accuracy of measuring instruments, appropriateness for function, and correctness of data are all determined through validation (Sanaky, 2021). The score obtained from a validity test will be used as a guide for choosing test items

(Muzaffar, 2016). Both empirical and logical validity were measured in this study. According to Riyani et al. (2017), empirical validity is founded on actual data or research experience, whereas logical validity is a test based on professional judgment or reasoning. Based on the findings, three experts—lecturers and a teacher—performed an expert assessment to determine logistical validity.

The instrument questions must be approved by expert judgment before the implementation of the first test and real test. This occurs by giving each question feedback and modifying the questions in response to that input. Three experts participated in the validity based on expert judgment: one science teacher and two biology lecturers. The Aiken Index is used to determine the validity of an instrument for expert judgment. Three expert judgments that were analysed using Aiken's V formula to evaluate the content validity of this study (Azwar, 2015). This is Aiken's V formula:

$$V = \frac{\Sigma s}{n(C-1)}$$

With the description,

V = the agreement index from the experts' rate

Σs = the average of the score gained from the rater minus the lowest score can be chosen

n = the total or number of experts

C = the number of categories can be chosen.

Next, the V Aiken coefficient is calculated. Criteria need to be defined in order to determine the validity of a test item. According to Irawan and Wilujeng (2020), the following criteria are used to assess the validity of content: below 0.4, the validity is invalid; between 0.4 and 0.8, the validity is moderate; and beyond 0.8, the validity is high. Table 3.4 shows the validity requirements based on the Aiken Index.

Table 3.4 Criteria for Validity using the Aiken Index

Index Validity (V)	Interpretation
$V > 0.8$	High Validity
$0.4 < V \leq 0.8$	Moderate Validity
$V \leq 0.4$	Invalid

(Irawan & Wilujeng, 2020)

Subsequently, each item is categorized using the V index that is obtained. This categorization gives specific categories for the instrument's content validity, which helps to understand the overall quality of the instrument, based on the level of agreement among expert raters (Irawan & Wilujeng, 2020).

The content-validity coefficient is determined using Aiken's formula, which is based on the opinions of expert panels including n individuals to a certain system. The degree to which the items reflect the measured construct is the basis for this (Hendryadi, 2017). Aiken introduced the idea of content validity, this information is evident from the validity standard, which is impacted by the rating scale and number of raters. Researchers can choose how many rating categories they wish to use to demonstrate the validity of the information. Aiken's content validity criterion is impacted by the quantity of rating categories (Anggraini et al, 2020). The result of the Aiken Test by expert judgment shown in Table 3.5, for the detailed result also shown in Appendix B.3.

Table 3.5 The Result of Aiken Test

Test Item	Tier 1		Tier 3	
	Aiken's Index (V)	**Interpretation**	**Aiken's Index (V)**	**Interpretation**
1	1.0	High	1.0	High
2	1.0	High	1.0	High
3	1.0	High	1.0	High
4	1.0	High	1.0	High
5	1.0	High	1.0	High
6	1.0	High	1.0	High
7	1.0	High	1.0	High
8	0	Invalid	0.67	Moderate
9	0.3	Invalid	0.67	Moderate
10	1.0	High	1.0	High
11	1.0	High	1.0	High
12	1.0	High	1.0	High
13	1.0	High	1.0	High
14	1.0	High	1.0	High
15	0.3	Invalid	1.0	High
16	0	Invalid	1.0	High
17	0.67	Moderate	1.0	High
18	1.0	High	1.0	High
19	1.0	High	1.0	High
20	1.0	High	1.0	High
21	1.0	High	1.0	High
22	1.0	High	1.0	High
23	1.0	High	1.0	High
24	0.67	Moderate	1.0	High

This table revealed that certain test items imply moderate scores. This outcome demonstrates that there is a strong connection between scores and shows that it can be used as a valid measuring tool (Slamet & Wahyuningsih, 2022). Appendix C.2 contains the expert judgment forms completed by the three experts.

The expert judgments examined whether the questions adequately cover the key concepts and content areas of biology education and whether they effectively capture the range of knowledge and misconceptions that students may have. The experts provided valuable feedback and insights on the clarity, accuracy, and overall quality of the questions. They assessed the alignment between the questions and the targeted content, ensuring that the questionnaire effectively measures the intended construct of students' conceptions and misconceptions in biology education. By involving experts in the field of biology education, the researchers can benefit from their expertise and judgment to enhance the content validity of the instrument. The feedback provided by the experts guided the refinement and improvement of the questionnaire, ensuring that it is a valid and reliable tool for assessing students' understanding and misconceptions in the context of the research.

By comparing the value of the R-table with the R-count (corrected item-total correlation), one can determine the validity of the instrument to be utilized as a measuring tool. When the R-count value is greater than R-table value, the test items are said to be valid. An R-count value lower than the R-table value means the test item is invalid. The table of correlation coefficient from data analysis with significant values is Table R (Indrawan & Yaniawati, 2014). Additionally, to determine the degrees of freedom, we can use the formula df=N-2 where N is the number of samples used. For example, if the number of samples is 34, then the degrees of freedom become 32. The degress of freedom value is used to find the RTable value corresponding to the 5% significance level.

A correlation test was used to evaluate validity and appropriateness of the questions. Without looking at the impacted variables, this test looks at the association between two variables (Cahyaning, 2017). Students in grade 8 were given the questions.

One class (34 students) participated in the validation test. The appropriateness, relevance, and clarity were evaluated by an analysis of the

responses, and the questions were refined and modified by the researcher as needed in light of the validation phase's results.

Thirteen of the twenty questions that were examined during the initial testing phase, which had 34 respondents, were deemed valid. These questions are Q2, Q3, Q6, Q7, Q8, Q9, Q11, Q12, Q14, Q15, Q17, Q18, and Q19. This indicates that the 13 questions are valid and can be utilized straight to the real test because they satisfy the validity requirements. The remaining seven questions (Q1, Q4, Q5, Q10, Q13, Q16, and Q20) were not included in the real test. By adding more details regarding the difficulties students confront, the questions were updated to increase their validity. Table 3.6 details the findings of the validity test that employed the student-held questions shown in Appendix B.1.

Table 3.6 Content Validity Test Results

Question	Tier	RCount	RTable (5%)	Interpretation	Decision
1	1	0.283	0.338	Not Valid	Not Used
	3	0.424	0.338	Valid	
2	1	0.418	0.338	Valid	Used
	3	0.625	0.338	Valid	
3	1	0.357	0.338	Valid	Used
	3	0.517	0.338	Valid	
4	1	0.242	0.338	Not Valid	Not Used
	3	0.505	0.338	Valid	
5	1	0.210	0.338	Not Valid	Not Used
	3	0.620	0.338	Valid	
6	1	0.672	0.338	Valid	Used
	3	0.419	0.338	Valid	
7	1	0.359	0.338	Valid	Used
	3	0.767	0.338	Valid	
8	1	0.647	0.338	Valid	Used
	3	0.458	0.338	Valid	
9	1	0.647	0.338	Valid	Used
	3	0.430	0.338	Valid	
10	1	0.129	0.338	Not Valid	Not Used
	3	0.324	0.338	Not Valid	

Question	Tier	RCount	RTable (5%)	Interpretation	Decision
11	1	0.665	0.338	Valid	Used
	3	0.529	0.338	Valid	
12	1	0.490	0.338	Valid	Used
	3	0.491	0.338	Valid	
13	1	0.554	0.338	Valid	Not Used
	3	0.277	0.338	Not Valid	
14	1	0.352	0.338	Valid	Used
	3	0.767	0.338	Valid	
15	1	0.599	0.338	Valid	Used
	3	0.604	0.338	Valid	
16	1	0.498	0.338	Valid	Not Used
	3	0.255	0.338	Not Valid	
17	1	0.567	0.338	Valid	Used
	3	0.536	0.338	Valid	
18	1	0.733	0.338	Valid	Used
	3	0.569	0.338	Valid	
19	1	0.594	0.338	Valid	Used
	3	0.419	0.338	Valid	
20	1	0.647	0.338	Valid	Not Used
	3	0.318	0.338	Not Valid	

3.3.3 Reliability Test

Before doing the actual research, a reliability test was completed. Reliability is the degree to which a measurement is consistent. A measurement result can be trusted if it is conducted on the same group of participants multiple times and yields relatively consistent results (Matondang, 2009). A popular technique for evaluating a scale or questionnaire's internal consistency is reliability testing with Cronbach's alpha (Creswell, 2012). The reliability interpretation shown in Table 3.7.

Table 3.7 Reliability Interpretation

Correlation coefficient	Reliability criteria
0.80 – 1.00	Very High
0.60 – 0.79	High
0.40 – 0.59	Moderate
0.20 – 0.39	Low
0.00 – 0.19	Very Low

All 13 valid questions from the questionnaire were included in this test, along with the appropriate answers from the students. The replies were converted into binary numbers, with "0" denoting erroneous or poorly thought-out responses and "1" denoting accurate responses. This reliability test measures students' conceptual comprehension and misconceptions by evaluating the consistency and stability of the four-tier diagnostic test. The result of the reliability test shown in Table 3.8.

Table 3.8 The reliability test result

	Cronbach's Alpha	N of Items
1st Tier Question	0.869	13
3rd Tier Question	0.864	13

Only valid questions are used for reliability testing. This is in line with the questions' stability and consistency, which are governed by question validity criteria; if the questions are invalid, their consistency is said to be reduced. At tier 1, the reliability value was 0.869, while at tier 3, it was 0.864. The reliability results are shown in detail in Appendix B.2. According to Riyani et al. (2017), instruments in these two categories are considered to have very high correlation coefficient values because they range between $0.80<n\leq1.00$. Therefore, it can be claimed that this instrument's high question consistency makes it appropriate for application in actual research as a tool to evaluate or diagnose cell related misconceptions.

The reliability coefficient was ascertained by the researcher using statistical methods in SPSS, including Cronbach's Alpha. When a test item has a high reliability coefficient, it means that the item can be used for the real test. The researcher can ensure that the results of the data collected from the real test are reliable and can see how much understanding as well as misconception students have on the topic of cell.

3.4 Research Procedure

The researcher divided the steps into three stages in order to finish this study: the preparation stage, which is done before the actual research being conducted; the implementation stage, which uses the actual instrument; and the completion stage, which is used to finish the data based on the outcomes of the implementation stage. Those three stages are explained as follows.

1. Preparation Stage
   a. Identify a research problem: At this point, the researcher selects "Misconception in Junior High School Students' Understanding of Cell" as the study problem.
   b. Analyze the topic of Cells contained in Science Textbook 8th Grade in Merdeka Curriculum to be developed in preliminary study.
   c. Perform a thorough analysis of the literature on the subject to learn more about earlier research and conclusions about misconceptions in biology teaching.
   d. Make the open-ended questions in scientific knowledge and scientific reason form for preliminary study.

e. Before distributing the preliminary study instrument to students, validate the questions to the supervisor first.
f. Collect and generalize the results of preliminary study.
g. Make a four-tiered diagnostic test with questions about various aspects of cell topics. Make sure the questions address the highlighted misunderstandings and are in line with the study questions.
h. Validate and revise the questions based on feedback given by the supervisor. Then, revise the questions by 3 expert judgments.

2. Implementation stage
   a. Select a sample of 8th graders who have studied cell subjects from a junior high school in Bekasi, West Java.
   b. Contact the school and provide a letter of permission to conduct research from the campus, Appendix C.3 and Appendix C.4 provided the letters.
   c. To assess the construct validity and reliability of the instrument, administer the validation test using the first 20 questions.
   d. Based on the findings of the validation test, use statistical analysis to ascertain the construct validity and reliability of the instrument.
   e. To improve the validity and reliability of the instrument, make revisions based on the results of the validation test.
   f. Based on the findings of the content validity and reliability tests, determine which valid questions will be administered.
3. Completion stage
   a. Collect the students' answers from 13 valid questions and document the outcomes for additional examination.
   b. Analyse and report the data from the real test to find out how common misconceptions are among the students and how useful the instrument is overall.
   c. Have semi-structured interviews with a chosen of students to learn more about their conceptual grasp and misconceptions.
   d. Determine the misunderstandings held by the students and potential areas where biology education could be improved by interpreting the results of the data analysis and interviews.

e. Made the conclusions, implications, and recommendations based on the results of the data analysis.
f. Publishing the research findings in a regional journal.

From the planning stage through the final stage, the research employs a systematic approach to guarantee the validity and reliability of the instrument used to measure students' misconceptions. In addition to helping to validate the instrument, the utilization of several pilot tests, interviews, and statistical analysis offers insightful information about students' conceptual knowledge. The study's conclusions can benefit the field of science education by assisting teachers in creating practical plans that encourage students to understand biology in a meaningful way.

A thorough record of the study process and its approval is provided by the research documentation. The technique and results of this study are repeatable and can be thoroughly re-examined in a variety of different academic settings and geographical locations.

At the completion stage, proof of journal publication is attached in Appendix C.1. Then before submitting the thesis, a plagiarism check is carried out, proof of similarity index is attached in Appendix D.4.

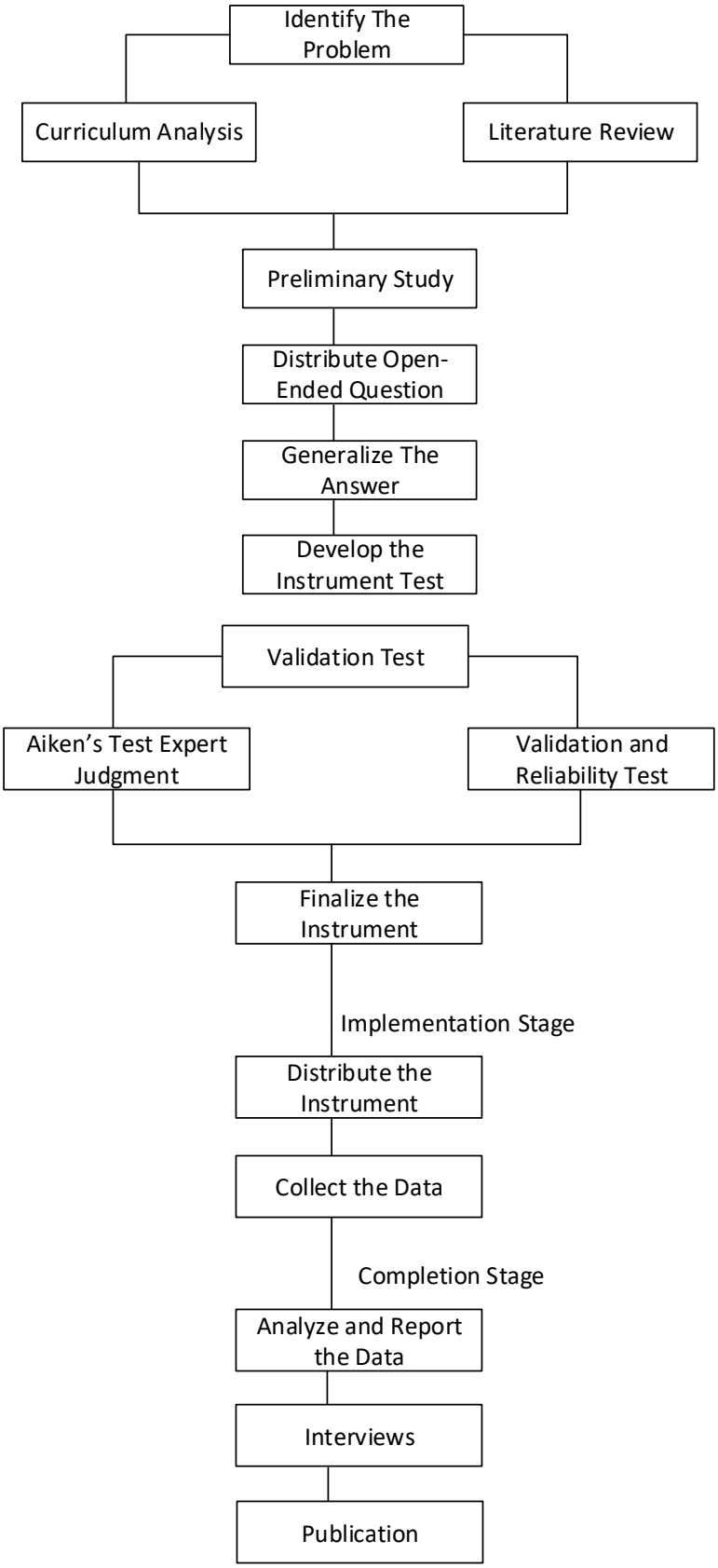


Figure 3.1 Stage of Research

3.5 Data Analysis

The classification of students' conceptions explained in Chapter 2 was used. The percentages for every category were then calculated. This computation was made with a simple formula that involved calculating the proportion of the results that fell into each group. The following formula is applied to these calculations:

$$P = \frac{s}{N}\ x\ 100\%$$

Description:

P = Percentage of each category

N = Number of the students for each group

s = Total number of students

As an illustration, suppose that 100 students took part in the study and that the following categories were applied to their responses:

20 students provided False Positive (FP) responses, 15 students provided False Negative (FN) answers, 5 students provided Misconceptions (M), 10 students provided Lack of Knowledge (LK), and 50 students displayed Scientific Knowledge (SK). The number of students in each category was divided by the total number of participants (100), and the result was multiplied by 100 to determine the percentage in each category. Consequently, the following percentages were ascertained: 20% for FP, 15% for FN, 5% for M, 10% for LK, and 50% for SK.

Based on the students' answers to the four-tier diagnostic test, these percentages show how the students' conceptions were distributed within each category. The researcher can proceed with additional analysis and interpretation of the data after calculating the percentages of misconceptions in each question. Based on this analysis, future teaching tactics or interventions may be guided by the identification of the particular concepts or themes that are producing the most misconceptions among students.