CHAPTER III
METHODOLOGY

3.1 Research Design

In this study, a survey research design was utilized. The survey design represents one of the procedures within the realm of quantitative research methods. The research design used in this research is cross sectional survey design. by distributing four tier test instrument Cross-sectional survey design is a research method that involves collecting data from a sample of individuals or groups at a specific point in time. data is collected from participants once, and various variables of interest are measured simultaneously. The purpose is to gather information about the prevalence, distribution, and relationships among variables at a specific time point. Common data collection methods in cross-sectional surveys include questionnaires (Cresswell, 2012).

3.2 Population and Sample

1. Population

   The population of this study consists of 9^{th} grade students from a specific junior high school in Bandung who have received instruction on the topic of plant reproduction in accordance with the 2013 curriculum.

2. Sample

   The sampling technique employed is convenience sampling (Cresswell, 2012).The sample for this research consists of 9^{th} grade students from one junior high school located in Bandung, West Java. The school selected a sample of three classes from which data will be collected for the research. The distribution of participants is presented in Table 3.1.

Table 3.1 Participant Distribution

Gender	Number of student	Percentage (%)
Male	53	52.5
Female	48	47.5
Total	101	100

3.3 Research Instrument

The instrument used in this study was a four-tier diagnostic test on plant reproduction. The researcher developed a test consisting of 30 questions (See Appendix A.1). specifically focused on the topic of plant reproduction. The researcher employed a validated instrument for data collection in this study, utilizing both student validation and expert judgment. This approach ensured the reliability and accuracy of the instrument used to measure students' misconceptions and understanding of the topic In addition to student validation, the researcher also sought three expert judgment. Experts in the field of plant reproduction or educational assessment provided their professional insights and feedback on the instrument. Their expertise helped ensure that the questions adequately captured the targeted misconceptions and aligned with the research objectives. By incorporating both student validation and expert judgment, the researcher strengthened the validity and robustness of the instrument used in the study (See Appendix A.2). This instrument underwent one stage of development, involving 32 students. In this stage, a total of 30 sets of questions were administered to the students through a Google Form. Subsequently, the results were analyzed for validity and reliability using SPSS. This rigorous validation process enhances the reliability of the data collected and increases the confidence in the findings related to students' misconceptions and understanding of plant reproduction. The allocation of questions can be observed in Table 3.2, where a total of 12 questions have been distributed.

Table 3.2 Question distribution

Sub topic	Concept	Question number
Vegetative reproduction	Natural vegetative	Q1
	Natural vegetative	Q2
	Natural vegetative	Q3
Plant pollination	Pollination process	Q4
	pollination type	Q5
	pollination type	Q6

Sub topic	Concept	Question number
Plant fertillization and seed dispersal	Seed dispersal	Q7
	Fertillization process	Q8
	Seed dispersal	Q9
	Fertillization peocess	Q10
Plant Reproduction technologies	Types of plant reproduction technology	Q11
	Types of plant reproduction technology	Q12

Based on Table 3.2 Each question in the instrument consists of four levels to assess students' understanding and confidence. Questions at level one are created in the form of multiple choice with four choices. The second level approximates the confidence level of answers at the first level, offering two options: "sure" and "not sure." The third level focuses on scientific reasons for the answers given in the first level, providing four options. The fourth level assesses the level of confidence in the answers provided at the third level. The sample question is shown in Table 3.3. Therefore, a total of 30 questions were tested and analyzed in this research.

Table 3.3 Question sample

Question 2	
Tier I	Pay attention to the following data! 1. Red Onion 2. Ginger 3. Potato 4. Galangal Based on these data the plants included in the rhizomes are? A. 1 and 3 B. 2 and 3 C. 1 and 4 D. 2 and 4
Tier 2	Are you sure with your answer? A. Sure B. Not sure

Question 2	
Tier 3	Why did you choose that answer? A. Because these plants have stems in the ground, these rhizomes are segmented and branched, in each book there are shoots that develop into new individuals B. because these plants have plant stems that spread over the ground. This rhizome has books, each book has shoots that can develop into new individuals. C. Because these plants have stems that experience swelling outside the ground. This rhizome has buds that can develop into new individuals. D. Because these plants have buds on the edges of the leaves, these buds can develop into new individuals.
Tier 4	Are you sure with your answer? A. Sure B. Not sure

In Table 3.3, the question aims to assess students' understanding of one of the characteristics of natural vegetative reproduction in rhizomes and provide an example. Students are then asked to identify the name of plants that possess rhizomes. Additionally, they are requested to select a reason and indicate their confidence levels for both the first and third levels. In this question, students are required to demonstrate their understanding of the concept of rhizomes, specifically Rhizomes are specialized underground stems found in certain plants. They are typically segmented and branched, with each segment capable of giving rise to new shoots that develop into new individuals. Rhizomes function as a means of asexual reproduction, enabling the plant to spread and establish new individuals from these underground stems.

As four -tier instrument will be tested by validity and reability test

1. Validity test

   Validity is a measure that indicates the extent to which an instrument accurately and appropriately assesses the content it intends to measure (Goodwin & Leech, 2003). It demonstrates the instrument's ability to effectively capture and represent the construct or concept it is designed to evaluate. In educational research, validity is crucial to ensure that the results obtained from the instrument are meaningful and relevant to the research

objectives. A valid instrument accurately reflects the specific knowledge, skills, or attitudes it aims to measure, providing reliable and trustworthy data for analysis and interpretation. Researchers employ various methods, such as content validity, construct validity, criterion validity, and face validity, to assess and establish the validity of their instruments in order to make meaningful inferences and draw accurate conclusions from the data collected.

In the context of content validity, the assessment was conducted with the assistance of three experts, among whom were a biologist and a science teacher this is a common approach to ensure that the content of the instrument aligns with the intended construct or concept. Content validity of an instrument entails an analysis of the problem representation's measurability, which can be established through expert consensus. This can be quantified using Aiken's Index, Aiken's Index, also known as Aiken's V or Aiken's Content Validity Coefficient, is a statistical measure used to quantify the content validity of an assessment instrument (Aiken, 1980). It's calculated using the following formula:

$$V = \frac{\sum s}{\mathrm{n(c-1)}}$$

where V is the item validity index; s is scores assigned by each rater minus the lowest score in the used category (s = r - lo, with r = rater category selection score and lo the lowest scores in the scoring category); n is the number of raters; and c is the number of categories that raters can choose. Derived from the computation of results using the V Index, The categorization of content validity within the instrument aligns with the values outlined in Table 3.4.

Table 3.4 Aiken's Indeks Validity Criteria

Aiken's Indeks	Validity Criteria
$0 \leq V \leq 0.4$	Low
$0.4 \leq V \leq 0.8$	Moderate
$0.8 \leq V \leq 1.0$	High

(Retnawati, 2016)

To observe the outcomes of the agreement among validators concerning the instrument, encompassing a total of 30 questions across diverse categories ranging from the lowest to the highest, refer to Table 3.5.

Table 3.5 Result of Aiken's Indeks

Question	Tier I		Tier III	
	Aiken's Indeks (V)	Validity Criteria	Aiken's Indeks (V)	Validity Criteria
1	0.8	High	0.8	High
2	0.8	High	0.8	High
3	0,8	High	0,8	High
4	1.0	High	1.0	High
5	1.0	High	1.0	High
6	0.3	Low	0.3	Low
7	0.7	Moderate	0.7	Moderate
8	1.0	High	1.0	High
9	0.7	Moderate	0.7	Moderate
10	1.0	High	1.0	High
11	1.0	High	1.0	High
12	1.0	High	1.0	High
13	0.7	Moderate	0.7	Moderate
14	0.7	Moderate	0.7	Moderate
15	1.0	High	1.0	High
16	1.0	High	1.0	High
17	0.8	High	0.8	High
18	0.7	Moderate	0.7	Moderate
19	1.0	High	1.0	High
20	0.7	Moderate	0.7	Moderate
21	07	Moderate	07	Moderate
22	1.0	High	1.0	High
23	0.7	Moderate	0.7	Moderate
24	0.7	Moderate	0.7	Moderate
25	0.7	Moderate	0.7	Moderate
26	0.7	Moderate	0.7	Moderate
27	0.7	Moderate	0.7	Moderate
28	1.0	High	1.0	High
29	0.5	Moderate	0.5	Moderate
30	1.0	High	1.0	High
Average	0.8	High	0.8	High

Referring to the data presented in Table 3.5, the outcomes of the content validity assessment across 30 questions indicate that the instrument is predominantly valid, falling within the moderate and high categories. A minor portion exhibits a low category. The average index for the Trebut instrument stands at 0.8, signifying a high category of validity. Based on the experts' feedback, the test was revised to incorporate their suggestions and recommendations. The revised test, which incorporated feedback from the experts, was then distributed to a student sample for a second round of testing. This testing phase involved administering the revised test to students from different schools. The purpose of the second student test is to assess the validity and reliability of the instrument, which was revised based on feedback from the first student test. This larger-scale testing aims to ensure that the instrument effectively measures student misconceptions about the topic of plant reproduction. The data collected from this student sample will contribute to the analysis and findings of the research. This iterative process of obtaining expert feedback, revising the test, and conducting student testing strengthens the validity, reliability, and generalizability of the instrument. It helps ensure that the instrument effectively measures the targeted misconceptions and aligns with the research objective. out of 30 questions, only 12 research instruments were used according to the validation results The student testing data for instrument validity yields 12 valid questions suitable for utilization, as evidenced in Table 3.6.

Table 3.6 Result of Validity test

Question	Tier	Test I			Test II		
		R count	Interpretation	Decision	R count	Interpretation	Decision
Q1	1	0.662	Valid	Accepted			
	3	0.595	Valid				
Q2	1	0.561	valid	Revision	0.432	valid	Accepted
	3	0.143	Not valid		0.340	valid	

Question	Tier	Test I			Test II		
		R count	Interpretation	Decision	R count	Interpretation	Decision
Q3	1	0.054	Not valid	Revision	0.342	valid	Accepted
	3	0.528	valid		0.498	valid	
Q4	1	0.476	Valid	Accepted			
	3	0.403	Valid				
Q5	1	0.350	Not valid	Revision	0.491	Valid	Accepted
	3	0.065	Not valid		0.474	Valid	
Q6	1	0.661	Valid	Revision	0.518	Valid	Accepted
	3	0.124	Not valid		0.293	Valid	
Q7	1	0.651	Valid	Accepted			
	3	0.383	Valid				
Q8	1	0.451	Valid	Accepted			
	3	0.584	Valid				
Q9	1	0.497	Valid	Accepted			
	3	0.514	Valid				
Q10	1	0.587	Valid	Accepted			
	3	0.483	Valid				
Q11	1	0.740	valid	Revision	0.293	Valid	Accepted
	3	0.042	Not valid		0.378	Valid	
Q12	1	0.665	Valid	Accepted			
	2	0.430	Valid				

Based on Table 3.6 the first test, a set of 30 questions was subjected to validation. After second validity test, 12 questions were considered valid and passed the validation criteria. These 12 questions were selected as the final instrument to analyze students' misconceptions on the topic of reproduction in plants (See appendix B.1 and B.2). Three different experts were consulted to obtain their opinions on the accuracy of the test items in relation to scientific knowledge, the appropriateness of the misconceptions being measured, the comprehensibility of the language used, and the overall suitability of the test for assessment purposes.

2. Reliability test

Reliability test is a measure used to assess the consistency and stability of an instrument or measurement over time. It indicates the degree to which the instrument produces consistent results when administered to the same group of individuals under similar conditions (Vehkalahti, 2000). The valid questions from the first and second tests conducted with students underwent two types of reliability tests. These questions demonstrated reliability with Cronbach Alpha scores. These scores indicate that the instrument has a reliable of internal consistency and can be considered acceptable for use (Brown, 2002). The specific Cronbach Alpha scores can be found in Table 3.7 total of 12 sets of questions were identified as suitable for diagnosing students' conceptions on reproduction in plants topics being studied.

Table 3.7 Result of reability test

Test	N of items	Cronbach's Alpha	
		Tier I	Tier III
I	7	0.615	0.658
II	5	0.634	0.619

3.4 Data analysis

To analyze the students' conception that occur based on combination students' answers in Table 3.8 The collected data were analyzed and the answers for each question based on five category were categorized into scientific

knowledge (SK), false positive (FP), false negative (FN), misconceptions (M), and lack of knowledge (LK). This categorization allows for a comprehensive assessment of students' responses, providing insights into their level of understanding, misconceptions, and areas where knowledge is lacking.. Scientific knowledge is categorized when students answer the first-tier correctly, express confidence in their first-tier response, answer the third-tier correctly, and exhibit confidence in their third-tier response. False positive refers to situations where students answer the first-tier correctly, express confidence in their first-tier response, but answer the third-tier incorrectly and are confident in their response. False negative occurs when students answer the first-tier incorrectly, express confidence in their first-tier response, but answer the third-tier correctly and are confident in their response. Misconception is categorized when students answer both the first-tier and third-tier incorrectly, but express confidence in both their first-tier and third-tier responses. Any combinations of responses that do not fall into the categories mentioned above are categorized as a lack of knowledge.

Table 3.8 Four Tier Diagnostic Test

	2nd Tier	**3rd Tier**	**4th Tier**	**Decision of Four-Tier Test**
True	Sure	True	Sure	SK
True	Sure	False	Sure	FP
False	Sure	True	Sure	FN
False	Sure	False	Sure	M
True	Sure	True	Not Sure	LK
True	Not Sure	True	Sure	LK
True	Not Sure	True	Not Sure	LK
True	Sure	False	Not Sure	LK
True	Not Sure	False	Sure	LK
True	Not Sure	False	Not Sure	LK
False	Sure	True	Not Sure	LK
False	Not Sure	True	Sure	LK
False	Not Sure	True	Not Sure	LK

The data analysis was performed using the Excel program, utilizing the combinations and decisions outlined in Table 3.8. In this analysis, correct answers for the first-tier and third-tier were scored as "1," while incorrect answers were scored as "0". For the second-tier and fourth-tier, the "sure" option was scored as "1" and the "not sure" option was scored as "0". Each category was analyzed on separate sheets within the Excel program. This approach allowed for a systematic analysis of the data, categorizing and scoring the responses according to the specific criteria outlined in the study.

Samples showing a score of "1-1-1-1" can be categorized as scientific knowledge (SK) since the students answered correctly in both the first-tier and third-tier questions, and expressed confidence in their responses for both tiers. For analyzing the lack of knowledge category, a code of "0" was assigned to the sequences corresponding to scientific knowledge (SK), false negative (FN), false positive (FP), and misconceptions (M). All other sequences that did not fall into these categories were considered as lack of knowledge and were coded as "1". This coding approach allowed for the differentiation and analysis of responses indicating a lack of knowledge separate from the other categories.

In the data analysis, frequencies and percentages were calculated to examine the distribution of responses within each category. Specifically, the percentages of scientific knowledge, false positive, false negative, and misconception were analyzed. This involved calculating the proportion of responses falling into each category relative to the total number of responses. These percentages provide insights into the prevalence and distribution of different types of responses and can help identify patterns or trends within the data.

To analyze the category of misconceptions that occur, It can be inferred that the percentage of misconceptions mentioned in Table 3.9 provides a reference for analyzing the occurrence of misconceptions. This table displays the specific percentages of responses categorized as misconceptions within the dataset. By examining the percentage of misconceptions, It allows for the assessment of the prevalence and frequency of these erroneous understandings among the participants. This information helps to identify the extent to which misconceptions

are present and provides valuable insights into the specific misconceptions that students may hold regarding the topic under study.

Table 3.9 Misconception category

Percentage	Category
0 -30 %	Low
31 – 60 %	Moderate
61 – 100 %	High

(Annisa et al., 2017)

3.5 Research Procedure

1. Prepation stage
   a. Literature study from any resource about misconceptions and four tier test
   b. Identifying research problem and research question
   c. Construct instrument

      Making 30 question four-tier test instument about plant reprdouction topic in junior high school
   d. validate the expertise, distribute the instrument questions to a group of experts in the field.
   e. Collect their feedback and consider their suggestions and comments.
   f. Initial validation of the instrument Distribute the questions in the form of Google Forms https://forms.gle/hTidtZ6LTbkioQcV7
   g. Running the validity test result
   h. Revision of instrument based feedback from expert
   i. Final validation of the instument that already revised based on validty and reability test on SPSS
   j. Running the validity and reability test.
2. Implementation stage

   Data collecting based on student answer of instrument by distributing test instrument junior high school students in Bandung the four-tier diagnostic test is administered to students with Google Forms

https://forms.gle/6gpmkmnnTeEJ43F19 (See Appendix D.1)The test consists of a total of 12 sets of questions. Students fill in the questions themselves, without any external assistance or guidance. There is only one questionnaire administered to the 9th graders. The link to the questions was administered on May 10th, 2023.

3. Completion stage
   a. Analysing the data that has been collected statistically. To diagnose students' conception, MS-Excel program is used
   b. Analyzing students' misconceptions based on data students' answer of instrument anlyzing the level students' understanding categories,
   c. Conducted interviews with 2 students during the Zoom meeting to find out the causes of misconceptions.
   d. Constructing the discussion based on data analysis and interpretation.
   e. Constructing conclusion and recommendation based on the result and discussion.
   f. The results of the data and discussion of the research after being guided by the two supervisors, then obtaining permission from the supervisor for "thesis defense" (see Appendix C.5)
   g. The research was approved by the supervising lecturer for conducting the thesis trial. Subsequently, the researcher submitted the research paper to SINTA 3 in the Untan Mathematics and Science Education Journal. (see Appendix C.6)

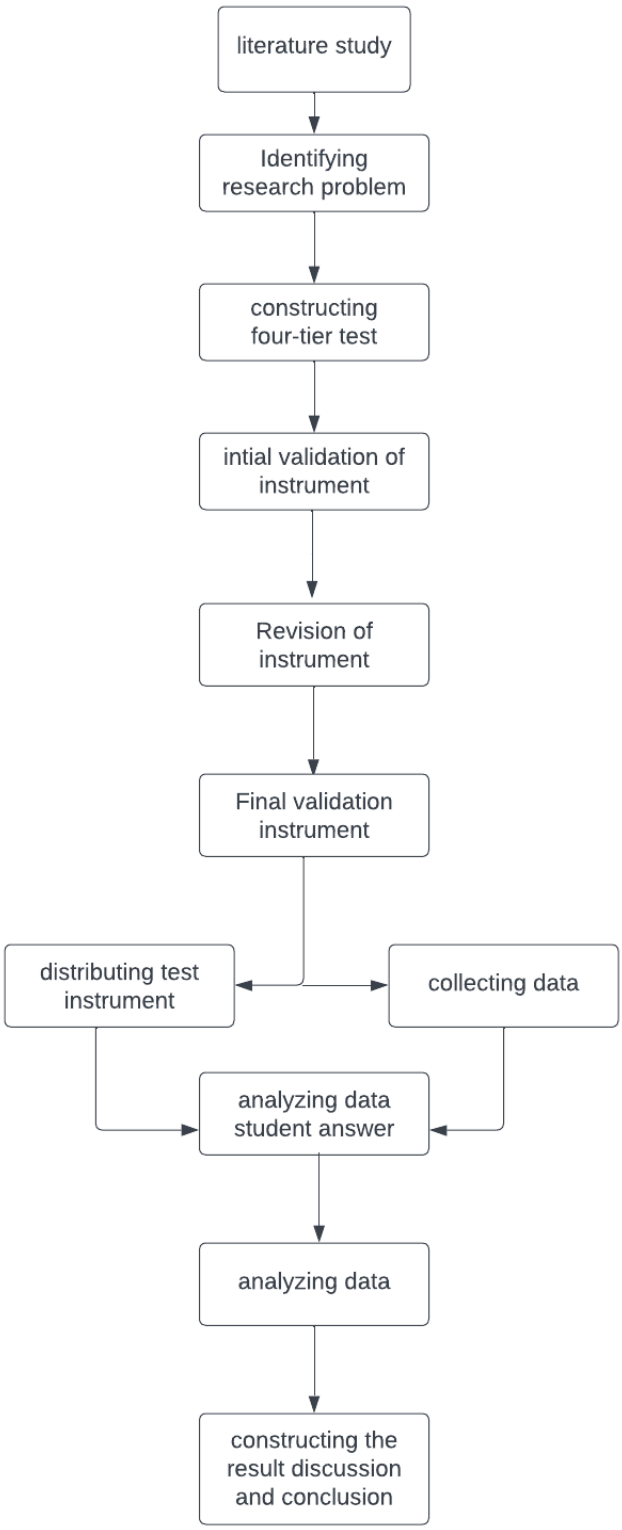

Figure 3.1 Research Procedure