Biosfer: Jurnal Pendidikan Biologi

Journal homepage: http://journal.unj.ac.id/unj/index.php/biosfer

Development of four-tier multiple-choice diagnostic instrument to identify middle school students' misconceptions about biology material

Tasya Novian Indah Sari[1]*, Yuni Pantiwati[2] , Ahmad Fauzi[3]

Biology Education, Faculty of Teacher Training and Education, Universitas Muhammadiyah Malang, Indonesia

*Corresponding author: tasyanovian@gmail.com

ARTICLE INFO

Article history
Received: 14 May 2023
Revised: 11 May 2024
Accepted: 16 May 2024

Keywords:
4TMC
Classificational misconceptions
Correlational misconceptions
Diagnostic evaluation
Theoretical misconceptions

ABSTRACT

Misconceptions are a condition where students have a mismatch between personal and scientific concepts that influence the learning process. Efforts to identify misconceptions using the Four-Tier Multiple Choice (4TMC) Diagnostic instrument. The aim of the research was to develop the 4TMC diagnostic instrument to identify misconceptions and types of misconceptions of Class VII students regarding the classification of living things. Type of research Research & Development. The research was carried out October 2021-March 2022 at one of the State Middle Schools in Magetan Regency, Indonesia. Subject 4TMC students 197 Class VII students. Data collection techniques using tests and non-tests. Non-test instruments in the research include teacher interview sheets and expert validation sheets. Meanwhile, the test instruments include the two-tier open ended and 4TMC tests. The research results showed that the assessment by experts and teachers obtained a feasible instrument. Testing the validity of the question items obtained 43 valid questions with a reliability degree of 0.848. The questions consist of 10 difficult questions and 33 medium questions. Furthermore, 6 different power questions were poor, 13 different power questions were sufficient and 24 different power questions were very good. Apart from that, the results of the misconception test showed that 47% of students had moderate misconceptions, 28% understood the concept, 17% did not understand the concept and 7% made errors. The types of misconceptions that occur among students include three forms, namely classificational misconceptions 44%, correlational misconceptions 33% and theoretical misconceptions 13%. The results of instrument development can be used again to identify misconceptions and types by teachers or future researchers.

Sari, T. N. I., Pantiwati, Y., & Fauzi, A. (2024). Development of four-tier multiple-choice diagnostic instrument to identify middle school students' misconceptions about biology material. *Biosfer: Jurnal Pendidikan Biologi, 17*(1), 257-273. https://doi.org/10.21009/biosferjpb.35289

INTRODUCTION

Science learning is learning that combines several concepts in one science study. The concepts incorporated in scientific learning are mathematics, physics, and biology (Noerinda, 2014). Science learning often causes difficulties for students. The material in science learning generally refers to facts, theories, concepts, and principles that must be studied and taught in science learning. The science learning process requires a complex reasoning process because science involves complex and contextual problems (Sa'diyah, 2023; Wulandari et al., 2023). If the learning carried out cannot provide visualization or a real picture for students, this provides an opportunity for misconceptions to occur. Kurtuluş & Tatar (2021) convey that science learning close to misconceptions in learning. One of the fields of science studies that is often frequented by misconceptions is the field of biology. Misconceptions often occur in the field of biology, because many concepts are related to one another (Andariana et al., 2020). Students who study biology material tend to memorize concepts in the material rather than understand the existing material. If the concept received by students is not appropriate, it is difficult to improve it again (Ramdani et al., 2021). This is what causes difficulties for students and often interpret the concept itself which results in misconceptions (Suprapto, 2020). If there is a misconception about one concept and it is permanent, it will affect other concepts. This can be a barrier to understanding further biological material (Madukubah & Taiyeb, 2018).

Students who experience misconceptions have errors in understanding material concepts, thus causing incompatibility between personal concepts and scientific concepts (Nurdiansyah et al., 2022). In learning activities, misconceptions are natural things in the process of forming the basis of knowledge (Suprapto, 2020). However, this cannot be allowed to continue because student misconceptions can affect the teaching and learning process in the classroom. Student misconceptions can also cause students' understanding of a concept to be inconsistent. Misconceptions that occur in students can affect the next learning process (Putri et al., 2022).

Identification of misconceptions is one of the first important steps to finding out the extent to which students understand concepts in material in learning activities (Nurul Humaidi et al., 2023). To find out students' misconceptions, a tool or instrument is needed that can be used to determine students' understanding of material (Syaifuddin et al., 2022). The instrument in question is using the Diagnostic Test instrument. One form of the diagnostic test is the Four-Tier Multiple Choice (4TMC) (Laliyo et al., 2021). The 4TMC diagnostic instrument is an advanced instrument or development of the Three-Tier Multiple Choice instrument (Caleon & Subramaniam, 2010). The 4TMC instrument is claimed to be the most effective instrument for solving students' misconceptions (Laliyo et al., 2021). This instrument was developed and designed to find out how strong students are in understanding the concepts they have learned through the level of student confidence in answering a question (Ramdani et al., 2019). This test can provide a more specific or clearer picture in grouping the level of students' understanding of the concept (Muzakki et al., 2023). In addition to knowing the level of students' understanding of concepts, this test can be used to determine the level of students' confidence in answer items and reasons (Caleon & Subramaniam, 2010). Therefore, this instrument can be used to detect misconceptions in students. The use of this instrument can assist teachers in measuring and identifying misconceptions in students (Sari et al., 2023). The use of the 4TMC diagnostic test can show the level of understanding of students when answering these questions and how students think (Laliyo et al., 2021). The results (Kaltakci-Gurel et al., 2017) showed that the 4TMC diagnostic test can be used to assess conceptual understanding or misunderstanding of certain material in a large group.

The results of an interview with the Science Teacher at Junior High School 1 Bendo Magetan show that the learning method used by the teacher recently during the pandemic was blended or offline online. Offline learning activities are carried out with limited face-to-face meetings by adjusting the rate of increase in COVID-19 in Magetan Regency and continued online learning via WhatsApp group. Not all students can study independently when face-to-face meetings are limited so that they learn biological concepts which are generally very complex. This difficulty affects students' thinking patterns, namely, students tend to memorize the concepts they learn from books or material provided, rather than understanding the concept of classification of living things.

Evaluation activities carried out during a pandemic could not run optimally. One of the evaluation activities carried out by the teacher is the daily repetition activity at the end of the lesson. Forms of daily test questions used by Class VII Teachers of Junior High School 1 Bendo, Indonesia on basic competency 3.2 Classifying living things based on the observed characteristics of living things in the form of multiple-

choice questions totaling ten questions. The questions used do not cover all the material taught in Basic Competency 3.2. Based on this, not all the material that students have studied is used as evaluation material.

For example, questions about the characteristics of living things are still limited to examples of living things without adding explanations regarding their characteristics. The limitations of these evaluation questions cause students to be limited in studying the classification of living things, grouping living things from general characteristics to special characteristics based on each living creature's special similarities and differences. It's followed by Elvianti's (2020) research, that many students still lack an understanding of the special characteristics of living things.

Facts in the field are also supported by research results by Mariyadi & WA (2023) that students still experience misconceptions in science learning amounting to 66.7% in physics material. Hasanti & Zulyusri (2021) revealed that among junior high school students, the highest misconceptions were also found in biology material. Findings by Andriani et al. (2021) the percentage of students' understanding of the concept of the misconception category regarding the concept of substance pressure was 43.125%. Meanwhile, it is also known that the percentage of students' misconceptions in the high category is 25%, the medium category is 62.5%, and the low category is 12.5%. Finally, the research results Syarif et al. (2023) that high school students also had 33.1% misconceptions about cell material. Some of the results of this research indicate that misconceptions in science subjects need to be considered and followed up.

The use of diagnostic tests to identify misconceptions has been carried out by several previous studies. Research conducted by Pakpahan et al. (2020) regarding the analysis of misconceptions on the nervous system material using the 4TMC instrument. Another research was conducted by Widiastutik & Isnawati (2021) regarding the profile of misconceptions about protein synthesis materials using the 4TMC instrument. Research conducted by (S. Wulandari et al., 2021) on the profile of misconceptions of genetic material using the 4TMC instrument. Meanwhile, research was conducted by Agustina & Indana (2022) regarding misconceptions about protist material using the 4TMC instrument. Research conducted by Afifah & Asri (2020) profiles misconceptions about the structure and function of cells using the 4TMC instrument. The last research on the development of the 4TMC instrument (Firdaus & Rahayu, 2020) identified misconceptions about carbohydrate catabolism.

There is research on the development of the 4TMC instrument to identify misconceptions about biological material, but the material for classifying living things for the development of the instrument is still in the form of a two-level diagnostic test. In addition, the development of the 4TMC instrument is still not equipped with a mapping of the types of misconceptions experienced by students. One of the studies that have mapped the types of misconceptions using the 4TMC instrument on high school student ecosystem material (Nurhidayah et al., 2020).

Based on the description that has been presented, the misconception is a serious problem in learning activities and requires handling as soon as possible. One of the efforts to handle misconceptions in students is to identify misconceptions. The most effective effort to identify misconceptions is to use diagnostic test instruments. The 4TMC instrument is a diagnostic instrument that has several advantages compared to previous instruments. Based on the description above, it is necessary to conduct research, namely the development of a 4TMC diagnostic instrument to identify misconceptions along with mapping the types of misconceptions in the classification of living things for Class VII Junior High School students.

METHODS

Research Design

Research on the development of a 4TMC diagnostic instrument was carried out using a development model (Caleon & Subramaniam, 2010) modified from Treagust (1988). The procedures for developing this instrument include (1) Defining the Content Boundaries of the Study, (2) Exploratory Phase, (3) Content Validation and Piloting, and (4) Construction, Validation, and Administration.

Population and Samples

The research was conducted up to the validation and trial stages. Research and development will be carried out from October 2021 to March 2022 at Junior High School 1 Bendo Magetan. The research subject involved two groups of students. 9 Class VIII students have taken material on the classification of living things as test subjects for the two-tier open-ended instrument to obtain alternative

misconceptions as material for the preparation of the third tier on 4TMC questions. Class VII students totaled 197 students who had taken material on the classification of living things for the 4TMC instrument trial.

Instrument

Collecting test data using the two-tier open-ended test and the 4TMC diagnostic test which totals 45 questions. A two-tier open-ended test was used to identify alternative misconceptions about the classification of living things. The 4TMC diagnostic test is used to find out how many misconceptions students have which will later be interpreted according to (Caleon & Subramaniam, 2010) provisions and types of misconceptions. Non-test data collection is in the form of teacher interview sheets for problem identification, material validation, assessment validation, and science teacher validation.

Procedures

1. Defining the Content Boundaries of the Study

This stage is the stage of compiling propositional questions and concept maps on living things classification material using living things classification textbooks and living things classification syllabus. The formulation of these questions will later be examined by a biology lecturer who concentrates on the classification of living things and one science teacher. The focus of the questions includes material on the classification of living things and the classification of animals and plants. The results at this stage are expressed as the product I (Caleon & Subramaniam, 2010).

2. Exploratory Phase

This stage is intended to obtain information about misconceptions in students. Proposal questions that have been prepared are made into instruments in the form of a two-tier open-ended. The purpose of this stage is to obtain alternative answers regarding the concept of the classification of living things. Students' answers on the two-tier open-ended test will be tabulated and analyzed which reasons are identified as misconceptions. Answers that identified misconceptions will be included in the reasons for the question of the third level of the 4TMC instrument. The results at this stage are expressed as product II (Caleon & Subramaniam, 2010).

3. Four-Tier Multiple Choice Test Design

The results of collecting alternative information on students' misconceptions using a two-tier open-ended instrument were developed into a 4TMC question which was placed at the third level of questions (reasons) with correct concept patterns and concepts with misconceptions. 4TMC questions consist of four multiple-choice questions in one number. The results at this stage are expressed as product III (Caleon & Subramaniam, 2010).

4. Content Validation and Piloting

The two-tier open-ended instrument that has been transformed into a 4TMC instrument will be validated by the assessment expert, material expert, and the science teacher concerned using the validation sheet provided. After the 4TMC instrument has been validated by experts and is feasible to be tested, the instrument will enter product trials. The product trial involved approximately 8 Classes of Class VII students of Junior High School 1 Bendo (Caleon & Subramaniam, 2010).

Data Analysis Techniques

The developed 4TMC diagnostic instrument was analyzed for validity. The validity data of the instrument was obtained from the validation of 3 validators, namely 1 science teacher at Junior High School 1 Bendo Magetan and 2 expert lecturers, namely experts on living things classification materials and experts on the assessment of biology education study programs at the University of Muhammadiyah Malang. The results of expert validation will be analyzed quantitatively using a Likert scale to determine the feasibility of the 4TMC instrument. The scale used in the assessment is the Likert scale (1) very poor, (2) not good, (3) good enough, (4) good, and (5) very good. The results obtained will be interpreted according to the percentage of instrument eligibility. Furthermore, to test the validity of the items carried out with product-moment correlation. A reliability test was performed with alpha Cronbach. Different power tests with the help of Microsoft Excel while the level of difficulty with SPSS software. The results of each test will be interpreted to determine the criteria and categories of the developed instrument (Table 1).

Table 1
Interpretation of 4TMC Instrument Test Results

Aspect	Criteria	Category
Instrument eligibility.	0-25 %	Not feasible
	>25-50%	Not worth it
	>50-75%	Worthy
	>75-100%	Very worth it
The validity of the items	r-count > r-table	Valid
	r-count < r-table	Invalid
Question reliability	0.00 – 0.19	The degree of reliability is very low
	0.20 – 0.39	Low degree of reliability
	0.40 – 0.59	Moderate degree of reliability
	0.60 – 0.79	High degree of reliability
	0.80 – 1.00	The degree of reliability is very high
Different power criteria questions	≥ 0,40	Discriminatory power is very good
	0.30 -0.39	Good discrimination
	0.20 – 0.29	Discriminatory power is sufficient
	≤ 0,19	Distinguishing power is not good
Difficulty level of questions	TK < 0.3	Hard
	0.3 ≤ TK ≤ 0.7	Currently
	TK > 0.7	Easy

The misconception test of the 4TMC test results was carried out after the instrument could be tested on a large group and interpreted based on the test results on answers, level of confidence in answers, reasons, and level of confidence in reasons (Table 2). The level of confidence uses a 6 Likert scale, namely (1) just guessing, (2) very unsure, (3) not sure, (4) sure, (5) very confident, and (6) very sure. The results of the interpretation of the 4TMC instrument will be continued to find the magnitude of misconceptions in students. The percentage of misconceptions will be categorized based on the results of the percentage of misconceptions, namely 0%-30% including low misconceptions, 31%-60% including moderate misconceptions, and 61%-100% including high misconceptions. Finally, the types of misconceptions will be interpreted with the types of misconceptions that are already available in the 4TMC question grid.

Table 2
Interpretation of 4TMC Test Results

Category	Answer	Level of Confidence in Answers	Reason	Level of Confidence in Reason
Misconceptions	Correct	<3	Incorrect	≥ 4
	Incorrect		Incorrect	
	Correct	≥ 4	Incorrect	
	Incorrect		Incorrect	
Lack of understanding concept	Correct	≤ 3	Correct	≤ 3
	Correct		Incorrect	
	Correct	≤ 3	Correct	≥ 4
	Correct	≥ 4	Correct	≤ 3
	Correct		Incorrect	
	Incorrect	≤ 3	Correct	≤ 3
	Incorrect		Incorrect	
	Incorrect	≥ 4	Correct	
	Incorrect		Incorrect	
Error	Incorrect	≤ 3	Correct	≥ 4
	Incorrect	≥ 4	Correct	
Understand concept	Correct	≥ 4	Correct	≥ 4

RESULTS AND DISCUSSION

Defining the Content Boundaries of the Study

The content definition stage is used to determine material indicators that will be used as material for instrument development. This stage begins with determining basic competence based on Permendikbud Number 37 of 2018. The basic competence used in the development of cognitive basic

competence instruments is "3.2 Classifying living things based on observed characteristics". basic competence classification of living things requires students to be able to classify living things based on the characteristics observed. This is the opinion Gultom (2019) in studying the classification of living things students must be able to classify living things based on their characteristics and describe the characteristics of living things while at the same time distinguishing living things from one another.

The question indicators that have been developed will be implemented in the form of proportional questions. In preparing the proportional questions, he was assisted by experts on the material and the Middle School Science Teacher. Material experts at the mentoring stage are carried out by a Biology Education lecturer at the University of Muhammadiyah Malang with an area of expertise in living things classification material. Assistance by material lecturers aims to make the questions arranged by scientific theory and avoid misunderstandings in questions (Caleon & Subramaniam, 2010).

The mentoring stage for the Science Teacher was carried out by a Science Teacher at Junior High School 1 Bendo Magetan who concentrated on biology. This assistance is intended so that the material prepared in the questions is by the material in the realm of the intended educational level (Caleon & Subramaniam, 2010). Of the 69 questions compiled, 24 questions were not by the material for classifying living things for junior high school students. So at this stage of the 69 questions compiled, 45 questions were following the material for Grade VII Junior High School students.

Exploratory Phase

Proposal questions that have gone through the mentoring stage have been implemented in the form of two-tier open-ended or two-tier open questions. The purpose of the two-tier open-ended test is to obtain information related to students' errors in the classification of living things (Lestari & Susantini, 2020). Error information was obtained from student reasons written by students at the second level of questions. The characteristics of the students' reasons will be taken into consideration in compiling the third tier for 4TMC questions. This was done with the hope that the developed 4TMC instrument would be able to properly reduce misconceptions (Caleon & Subramaniam, 2010). The two-tier open-ended test on the classification of living things consisted of 45 questions. The open two-tier trial phase was carried out by 9 class VIII students who had already taken the classification of living things with the criteria of three students in the low, medium, and high groups. The selection of students in the low, medium, and high groups is based on the teacher's recommendations. The trial was carried out for 2 days, namely Friday 4-5 February 2022 offline at the Science Laboratory of SMPN 1 Bendo Magetan.

Four-Tier Multiple Choice Diagnostic Test Design

The 4TMC instrument for the classification of living things in this study is divided into 5 major materials, namely characteristics of objects in the surrounding environment, characteristics of living creatures, classification of living things, animal classification, and plant classification (Table 3). The design of the 4TMC instrument for material classification of living things consists of four-level multiple-choice in one number of questions (Figure 1). The first level questions are coded as question A. The first-level questions are the usual multiple-choice tests for answers. Question A consists of 3 distractor answers and 1 correct answer. The second level test is in the form of confidence in answers at the first level with a scale of 1-6 which is coded question B. The third level test is in the form of reasons for answers at the first level which are coded question C or reasoning questions. The fourth level test is in the form of confidence in answers at the third level (reasoning) with a scale of 1-6 which is coded question D. Patterns of answers and student choices in one number will be interpreted to determine student understanding (Table 2)

Table 3
Student Concept Understanding Profile with the 4TMC Instrument

No	Material	Indicators for Classification of Living Creatures	Question Number
1.	Characteristics of objects in the surrounding environment	Group objects in the surrounding environment	1 & 2
2.	Characteristics of Living Creatures	Determine the characteristics of living things	3 & 4
3.	Classification of living things	Classify living things based on the division of five kingdoms	5-9
4.	Animal classification	Grouping vertebrate animal classes based on their characteristics	10-18
		Grouping vertebrate animals based on body temperature	16-17
		Grouping classes of invertebrate animals based on their characteristics	19-35
5.	Plant classification	Group plants based on their characteristics	36-45

Question A
Question A is a first-level question in the form of an ordinary multiple-choice test for answers consisting of three distractors and one correct answer. The number of choices can be adjusted according to the school level (junior high/high school).
A. .....
B. .....
C. .....
D. .....
Question B
Question B is a second-level question in the form of the level of student confidence in choosing an answer at the first level. The level of trust uses a Likert scale with a scale of 1-6.
1. just guess
2. very unsure
3. not sure
4. sure
5. very confident
6. very sure
Question C
Question C is a third-level question in the form of reasons for answers at the first level. The nature of the questions at the third level is a matter of reasoning. The third level questions consist of three misleading reasons and one correct reason. The number of choices can be adjusted according to the school level (junior high/high school).
A. .....
B. .....
C. .....
D. .....
Question D
Question D is a matter of the fourth level in the form of the level of student confidence in choosing reasons for questions at the third level. The level of trust uses a Likert scale with a scale of 1-6.
1. just guess
2. very unsure
3. not sure
4. sure
5. very confident
6. very sure

Figure 1. 4TMC Diagnostic Test Design

Content Validation and Piloting Phase

The instrument assessment stage in this research involved experts and practitioners. The experts involved in this research were two lecturers from Biology Education, University of Muhammadiyah Malang as well as practitioners from junior high school science teachers. The material expert lecturer in this research is a lecturer who teaches material on the classification of living things. The results of the material expert assessment showed a result of 98% and the 4TMC instrument developed received the "Very Feasible" criteria to be continued at the next development stage. Meanwhile, validation by assessment experts is carried out by one Learning Evaluation lecturer. Based on the results of the assessment carried out by the assessment expert, they received a total score of 65 out of a total score of 85. The results of the assessment by the assessment expert showed a result of 76% and the 4TMC instrument developed received the "Very Feasible" criteria to be continued at the next development

stage.

Next is the instrument assessment stage by practitioners. The practitioners involved in this research were science teachers from SMPN 1 Bendo, Indonesia. The science teacher assessment was carried out to find out whether the question indicators and questions developed in this research were appropriate for the junior high school level. Based on the results of the assessment carried out by the Science Teacher, they received a total score of 39 out of a total score of 40. The results of the Science Teacher assessment showed a result of 98% and the 4TMC instrument developed received the "Very Feasible" criteria to be continued at the next development stage.

Table 4
Results of Expert and Practitioner Assessment of the 4TMC Instrument

No	Expert and Practitioner	Percentage
1.	Expert on the classification of living things	98%
2.	Assessment or evaluation expert	76%
3.	Science teacher	98%
	∑%	**91%**

Four-Tier Multiple Choice Trial Phase

Testing the validity of the question items is carried out by correlating the score of each question item with the score of all the questions (Sumardi, 2020). In this study, researchers used the product moment correlation formula to determine the validity of the 4TMC question items. Of the 45 questions tested, there were 43 valid and 2 invalid questions, namely numbers 1 and 38 (Table 5). The basis for decision-making in the validity test is seen from the calculated r value with the r table value. If the r-count value is smaller than r-table then the question item is invalid (Faradillah et al., 2020). The r-table value in developing this instrument is 0.139 with a significance value of 0.05 or 5%. This can be obtained from the formula df=N-2. The N value comes from the number of respondents involved in this research, namely 197 students. So according to the df formula, the df in r-table is 195.

Table 5
4TMC Question Item Validity Test Results

Question Number	R table	R caunt	Interpretation
Question 1	0.139	-0.374*	Invalid
Question 2	0.139	0.289*	Valid
Question 3	0.139	0.316*	Valid
Question 4	0.139	0.270*	Valid
Question 5	0.139	0.159*	Valid
Question 6	0.139	0.405*	Valid
Question 7	0.139	0.371*	Valid
Question 8	0.139	0.243*	Valid
Question 9	0.139	0.254*	Valid
Question 10	0.139	0.263*	Valid
Question 11	0.139	0.279*	Valid
Question 12	0.139	0.436*	Valid
Question 13	0.139	0.427*	Valid
Question 14	0.139	0.417*	Valid
Question 15	0.139	0.476*	Valid
Question 16	0.139	0.512*	Valid
Question 17	0.139	0.152*	Valid
Question 18	0.139	0.205*	Valid
Question 19	0.139	0.546*	Valid
Question 20	0.139	0.175*	Valid
Question 21	0.139	0.174*	Valid

Question Number	R table	R caunt	Interpretation
Question 22	0.139	0.293*	Valid
Question 23	0.139	0.459*	Valid
Question 24	0.139	0.392*	Valid
Question 25	0.139	0.349*	Valid
Question 26	0.139	0.454*	Valid
Question 27	0.139	0.509*	Valid
Question 28	0.139	0.353*	Valid
Question 29	0.139	0.332*	Valid
Question 30	0.139	0.404*	Valid
Question 31	0.139	0.524*	Valid
Question 32	0.139	0.156*	Valid
Question 33	0.139	0.515*	Valid
Question 34	0.139	0.427*	Valid
Question 35	0.139	0.476*	Valid
Question 36	0.139	0.360*	Valid
Question 37	0.139	0.385*	Valid
Question 38	0.139	0.063*	Invalid
Question 39	0.139	0.368*	Valid
Question 40	0.139	0.451*	Valid
Question 41	0.139	0.412*	Valid
Question 42	0.139	0.528*	Valid
Question 43	0.139	0.315*	Valid
Question 44	0.139	0.504*	Valid
Question 45	0.139	0.289*	Valid

Question items that are valid in the validity test will then enter the question reliability test stage. Data from the reliability test of the 4TMC instrument shows that Cronbach's alpha value is 0.848 (Table 6). The Cronbach's alpha value of 0.848 indicates that the degree of reliability is very high.

Table 6
4TMC Question Item Reliability Results

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
0.84	0.84	43.00

The next analysis of the questions is the different power of the 4TMC questions. The results of the test for differentiating power of questions show that of the 43 questions analyzed there are 6 questions with poor differentiating power, 13 questions with sufficient differentiating power, and 24 questions with very good differentiating power (Table 7).

Table 7
4TMC Problem Different Power Test Results

Score intervals	Category	Question number	Number of questions	Percentage
≥0.40	Very good	7,12,13,14,15,16, 19,23,24,26,27,29,30,31,33,35 ,36,37,39,40,41,43,43, and 44	24	56%
0.30-0.39	Good	-	-	-
0.20	Enough	2,3,4,6,8,10,11,18,21,22,25,28 and 34	13	30%
≤0.19	Not good	5,9,17,20,32 and 45	6	14%

The difficulty level of the question is the student's opportunity to answer the correct question. The results of the 4TMC difficulty level test showed that there were 10 questions in the difficult category and 33 questions in the medium category. Most of the diagnostic tests developed have a moderate level of difficulty, namely 77%, and difficulty with a percentage of 23%. A good diagnostic test is a diagnostic test that has a moderate level of difficulty.

Table 8
4TMC Difficulty Level Test Results

Difficulty level criteria	Category	Question number	Number of questions	Percentage
Difficulty level < 0,3	Hard	5,6,8,9,17,20,22,25,44 and 45	10	23%
0,3 ≤ difficulty level ≤ 0,7	Currently	2,3,4,7,19,11,12,13,14,15,16,18,19, 21,23,24,16,27,28,29,30,31,32,33,34,35,36,37,39,40,41,42,43,	33	77%
Difficulty level > 0,7	Easy	-	-	-

Misconception Test

The results of the 4TMC test for the classification of living things used in the study amounted to 43 questions which were declared valid and reliable. The test was attended by 197 Class VII junior high school students. The results of the 4TMC test carried out by students will be analyzed and interpreted to find out students' misconceptions. The results of the students' 4TMC answers will be categorized into understanding concepts, lack of understanding of concepts, misconceptions, and errors (Table 2). The results of the analysis of students' understanding of concepts show that 47% of students have misconceptions, 28% of students understand concepts, 17% of students do not understand concepts, and 7% of students experience errors (Figure 2).

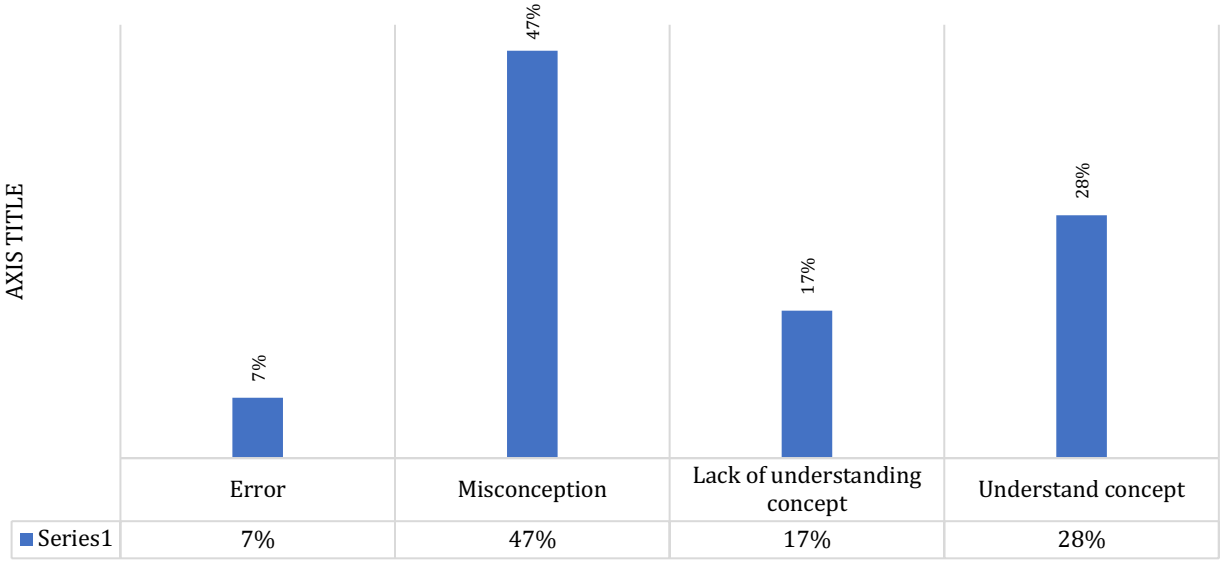

Figure 2. Student Concept Understanding Profile with the 4TMC Instrument

Table 9 shows the results of the percentage of misconceptions in each living creature classification material. Of the 5 materials presented, the animal classification material obtained the highest misconception value, namely 49%. Apart from the classification of animals, the material with the highest percentage of misconceptions is the classification of plants at 48% and the classification of living things in general, namely 48%. Finally, the material that obtained the smallest percentage of misconceptions was the characteristics of living things.

Table 9
Percentage of Misconceptions in the Material

No	Material	Question Number	Error	Lack of understanding concept	Miconception	Understand the concept
1.	Characteristics of living things	2,3,4	6%	10%	29%	55%
2.	Classification of living things	5,6,7,8,9	12%	22	48%	18%
3.	Animal classification	10,11,12,13,14,15, 16,17,18,19,20,21, 22,23,24,25,26,27, 28,29,30,31,32,33, 34,35	7%	17	49%	27%
4.	Plant classification	36,37,39,40,41,42, 43,44,45	5%	17	48%	29%

The results of the analysis regarding the types of misconceptions were carried out by interpreting the number of questions that had misconceptions with the types of misconceptions in the grid table. The results show that the misconceptions that occur among students cover three forms. The misconceptions that most often arise in this research are classificational misconceptions, correlational misconceptions, and theoretical misconceptions. The percentage of classificational misconceptions was 44%, correlational misconceptions 33%, and theoretical misconceptions 13% (Table 10).

Table 10
Misconception Type Test Results

Kind of misconception	Question number	Number of questions	Percentage of misconceptions
Theoretical Misconceptions	4,27,28,29,30,31,36,37,41,42	10	23%
The correlational misconception	6,7,8,9,22,23,24,26,32,33,34,35,39,40	14	33%
Classificational misconceptions	2,3,5,10,11,12,13,14,15,16,17,18,19,20 ,21,25,43,33,45	19	44%

A diagnostic test is a test that teachers can use to diagnose student misconceptions (R. E. Putri & Subekti, 2021). A good diagnostic test is a test that can provide an accurate picture of the misconceptions experienced by students (Fariyani et al., 2015). The 4TMC diagnostic test is a complementary test to the previous diagnostic test, because this test can convert confidence ratings at the first and third levels (Soeharto et al., 2019). The level of confidence in the 4TMC test can be used to differentiate the level of confidence in answers and the level of confidence in the reasons chosen by students can be used to identify the strength of students' understanding of concepts further. Apart from that, this test can be used to find out which material has more understanding and teachers can plan better learning to help reduce student misconceptions (Fariyani et al., 2015). Using tiers of reasons and level of confidence in the instrument can minimize students' guessing answers (Aprilanti et al., 2016).

Analysis of the Quality of 4TMC Instrument Items

The validation stage is carried out to determine the quality of the questions that have been created (Yasthophi & Ritongga, 2019). Validation is carried out to determine whether the instrument being developed is appropriate and can measure what is to be measured, in this development there are misconceptions among students (Mubarak et al., 2016). Instrument validation must be carried out by people who are experts in their field (Fariyani et al., 2015). The validation process in the development of the 4TMC instrument was carried out by assessment experts, material experts on the classification of living things, and junior high school science teachers (Caleon & Subramaniam, 2010). Material expert validation aims to ensure the validity of the content of the development evaluation instrument (Nofiana et al., 2014). This stage also aims to evaluate the evaluation product in terms of material for the classification of living things, including completeness of the material, suitability of the material to the level of education, and systematics of the material in the question instrument (Fahrurrozi & Laili

Rahmawati, 2021).

An instrument is said to be good as a measuring tool if it meets the test instrument requirements. These requirements include validity, reliability, distinguishability, and level of difficulty. One of the requirements for a good instrument is that it must be valid (Magdalena et al., 2021). The results of the validity test of the questions showed that there were 43 valid questions and 2 invalid questions, namely number 1 regarding the characteristics of living objects in the surrounding environment and number 38 regarding the characteristics of liverworts. Invalid questions cannot be included in further development of the 4TMC instrument. Even though there are invalid questions, there are representative indicators in the instrument developed.

The validity of the test items shows that the 43 test items developed can carry out the measurement function well. Apart from validity, good questions are supported by reliability. The reliability results of the 4TMC instrument only provide one value for all instruments. The instrument reliability value is 0.848. Question reliability testing is carried out on question items that have valid criteria. This shows that the validity of the test items is related to reliability. The more valid questions, the greater the reliability value of an instrument (Nurhalimah et al., 2022).

The results of the differential power test on the 4TMC instrument showed that there were 6 questions with poor differential power, 13 questions with fairly good differential power and 24 questions with very good differential power (Table 7). Questions with poor differentiation include examples and characteristics of animal-like protists, characteristics of fungi, groups of poikilothermic animals, characteristics of animals in the phylum Porifera, characteristics of animals in the Crustaceae class, and division of plants based on structure and morphology. Differential power is greatly influenced by the abilities of the group of students taking the test. If groups of students have different abilities, they can show a good differentiation index. So test group A and test group B will produce different power indices (Muluki & Bundu, 2020).

The 4TMC instrument's difficulty level test contained 33 questions with a moderate level of difficulty and 10 questions with a difficult level of difficulty (Table 8). Difficult questions include examples of indicators and characteristics of animal-like protists, plant-like protists, monera, fungi, poikilothermic animals, phylum porifera, phylum crinoidea, phylum coelentarata, monocotyledonous plants and the division of plants based on structure and morphology. Few students answered the questions correctly so the difficulty level index was low and the questions were difficult. While the questions are easy, many students answer the questions correctly, so the difficulty level index gets bigger and the questions get easier (Magdalena et al., 2021).

The level of difficulty and distinguishing power of questions is greatly influenced by the abilities of the group of students taking the test. But there is no relationship between the level of difficulty and the power of the different questions. Based on the analysis results in Table 7 and Table 8, questions in the difficult category may have poor or very good differentiating power. Items in the medium category can have poor, quite good or very good differentiating power (Muluki & Bundu, 2020).

The results of the level of difficulty and differentiation are influenced by the student's condition, so item validity and reliability are not influenced by the student's condition. The validity of a question item does not depend on the number of students who answer correctly. This is because questions that are answered correctly or incorrectly do not necessarily produce a high validity value. On the other hand, at the level of difficulty of the question, the greater the number of correct answers, the more the question item is classified as easy (Nurhalimah et al., 2022).

Misconceptions and Types of Misconceptions of Middle School Students

The research results showed that out of 197 students, 47% of students experienced misconceptions regarding the material on the classification of living things. This percentage is the highest percentage compared to other concept understanding profiles (Figure 2). This is in line with research (Gultom, 2019) Misconceptions in the material for classifying living things using a two-dimensional test show that students still experience misconceptions in classifying living things based on their characteristics. Results of misconception analysis by Fitri et al. (2018) using open multiple choice shows that Class VII students still experience misconceptions about the characteristics of living things.

In general, misconceptions in the field of biology are caused by several factors. In the field of biology, there are many interrelated concepts that students must master and the use of foreign terms in biology can affect students' understanding (R. Agustina et al., 2016; Puspitasari et al., 2019). The field of

biology is known as a field that has a lot of memorization, is complex and the process cannot be seen directly and has many theories (Aprilanti et al., 2016). When studying biology, students tend to memorize concepts rather than understand the concepts. This tendency causes different interpretations in students' understanding when it is lacking.

The material for classifying living things is quite complicated and complex (Kameswari, 2022). Students are required to be able to group living things based on their characteristics (Prihatiningsih et al., 2016). Apart from grouping students, they are also required to be able to describe the different characteristics of living objects, non-living objects, and groups of living things on Earth based on general to specific characteristics according to classification principles (Nurfajrianti, 2021). Material on the classification of living things uses a lot of Latin and foreign terms that are difficult for students to understand (Amalia, 2021). The use of quite a lot of foreign and Latin terms is one of the causes of the increasingly complicated classification of living things (Kameswari, 2022). The use of foreign terms is found in grouping living things based on class, phylum, and species. Apart from being complicated, the use of foreign terms which are quite complex can open up opportunities for misconceptions among students (Puspitasari et al., 2019). Misconceptions occur when students' understanding does not match scientific concepts (Amalia, 2021).

There are four materials developed in making indicators regarding the classification of living things, namely the characteristics of living things, the classification of living things, the classification of animals, and the classification of plants. Of the four materials developed, the animal classification material had the largest percentage of myconception, namely 49% (Table 9). Animal classification material requires students to be able to analyze the species of each animal. This aims to enable students to classify animals based on phylum or class characteristics. Animal classification material has complex characteristics from each phylum and class as well as many scientific names, making it difficult for students to understand the material (Yusriya et al., 2014). This difficulty will create students' understanding and create gaps for students to have misconceptions (Prianidya et al., 2015).

During the COVID-19 pandemic, learning activities at SMPN 1 Bendo Magetan, Indonesia were carried out online and blended. Online learning is carried out via the Whatsapp group by making attendance and sending material files. Online learning allows students to experience misunderstanding of the material presented by the teacher. So this causes misconceptions about the material presented (Noor, 2021). The results of research conducted by Ekawisudawati et al. (2021) The online learning methods used by teachers are not well understood by students. Learning methods like this make students less active in the learning process. Every student has different cognitive abilities. So not all students are compatible with the learning model used by the teacher and this is what influences students' understanding of concepts.

Apart from being used to identify misconceptions in students, the instrument can be used as a tool to determine the types of misconceptions. In general, the types of misconceptions are divided into three, namely correlational, classificational and theoretical misconceptions (Table 10). The type of student misconception can be identified by interpreting the question grid of the diagnostic instrument being developed. Each 4TMC question indicator has been mapped according to the type of misconception. The basis for this mapping is carried out by looking at the characteristics of the questions being developed along with the characteristics of the types of misconceptions. Apart from that, the development of the 4TMC diagnostic instrument which is equipped with mapping the types of misconceptions among students is still minimal, especially in the field of biology.

The results of the research show that the biggest type of student misconception in the material on the classification of living things is a classificational misconception (Table 10). Classification misconceptions are students' mistakes in classifying and grouping facts in an organized system (Fitriani & Rohaeti, 2020; Intan & Masriyah, 2020; Nurhidayah et al., 2020). In the material on the classification of living things, classification misconceptions relate to the grouping of living things based on the characteristics or characteristics of a group of living things. These results are following interviews with science teachers who stated that Class VII students were still lacking in grouping living things based on their characteristics. Research conducted by Gultom (2019) and Samiha et al. (2017) Regarding misconceptions in the material on the classification of living things, shows that students still experience misconceptions in grouping living things based on the characteristics or attributes of each group.

CONCLUSION

The 4TMC diagnostic test developed in this study produces a diagnostic instrument that is used to identify misconceptions and types of misconceptions of Class VII Junior High School students on the classification of living things. Good test questions are valid and reliable. In addition, the questions must also have different power and the appropriate level of difficulty. The results of the development in this study obtained the results of the questions and types of misconceptions, 4TMC questions, introduction to misconceptions along with 4TMC and school guidelines. There are 43 valid questions from the 45 questions developed. 43 valid questions have very high reliability, namely 0.848. There are 10 questions in the difficult category and 33 questions in the moderate category. Finally, the difference power test showed that there were 6 questions with poor different power, 13 questions with sufficient different power, and 24 questions with very good different power. The results of the misconception test showed that 47% of students experienced misconceptions in the moderate category, 28% of students understood the concept, 17% of students did not understand the concept and 7% of students experienced errors. Misconceptions that occur in students include three forms, namely classificational misconceptions of 44%, correlational misconceptions of 33%, and theoretical misconceptions of 13%.

The 4TMC diagnostic instrument on living things classification material can be used by teachers to identify misconceptions and types of misconceptions of Class VII students. In addition to the teacher, the instrument can be used by future researchers who take topics regarding misconceptions and types of misconceptions about the classification of living things and can be used in more than one school. The last research conducted has not gone further into the causes of misconceptions and types of misconceptions in students. So that in future research, you can add interview instrument sheets to students to reach the causes of misconceptions in students. So that further research can identify the misconceptions and types of student misconceptions in each item.

REFERENCES

Afifah, Y. N., & Asri, M. T. (2020). Profil miskonsepsi pada submateri struktur dan fungsi sel menggunakan four tier test. *BioEdu: Berkala Ilmiah Pendidikan Biologi*, *9*(3), 390–396. https://doi.org/10.26740/bioedu.v9n3.p390-396

Agustina, L., & Indana, S. (2022). Profil miskonsepsi siswa pada materi protista menggunakan four tier test. *BioEdu*, *11*(1), 60–67. https://doi.org/10.26740/bioedu.v11n1.p60-67

Agustina, R., Sipahutar, H., & Harahap, F. (2016). Analisis miskonsepsi pada buku ajar biologi sma kelas XII. *Jurnal Pendidikan Biologi*, *5*(2), 113–118. https://doi.org/10.24114/jpb.v5i2.4307

Amalia, N. 'Ilmi. (2021). Pengaruh penggunaan pbl berbantuan media interaktif powerpoint terhadap minat dan hasil belajar peserta didik materi klasifikasi makhluk hidup. *NCOINS: National Conference Of Islamic Natural Science (2021)*, 49–64. http://proceeding.iainkudus.ac.id/index.php/NCOINS/article/view/63

Andariana, A., Zubaidah, S., Mahanal, S., & Suarsini, E. (2020). Identification of biology students' misconceptions in human anatomy and physiology course through three-tier diagnostic test. *Journal for the Education of Gifted Young Scientists*, *8*(3), 1071–1085. https://doi.org/10.17478/JEGYS.752438

Andriani, D. W., Munawaroh, F., Qomaria, N., & Ahied, M. (2021). Profil miskonsepsi peserta didik berbasis taksonomi bloom revisi pada materi ipa konsep tekanan zat. *Jurnal Natural Science Educational Research*, *4*(1), 2654–4210. https://doi.org/10.21107/nser.v4i1.8400

Aprilanti, H., Qurbaniah, M., & Muldayanti, N. D. (2016). Identifikasi miskonsepsi siswa pada materi sistem ekskresi manusia kelas XI MIA SMA Negeri 4 Pontianak. *Jurnal Bioeducation*, *3*(2), 63–77. https://doi.org/10.29406/188

Caleon, I. S., & Subramaniam, R. (2010). Do students know What they know and what they don't know? Using a four-tier diagnostic test to assess the nature of students' alternative conceptions. *Research in Science Education*, *40*(3), 313–337. https://doi.org/10.1007/s11165-009-9122-4

Ekawisudawati, Wijaya, M., & Danial, M. (2021). Analisis miskonsepsi peserta didik pada materi asam basa menggunakan instrumen three-tier diagnostic test. *Chemistry Education Review*, *5*(1), 62–72. https://doi.org/10.26858/cer.v5i1.26359

Elvianti, M. (2020). Penerapan model discovery learning dapat meningkatkan aktivitas dan hasil belajar IPA pada materi ciri-ciri khusus makhluk hidup siswa kelas VI MIN 8 Aceh Barat Daya pelajaran 2018/2019. *Jurnal Pendidikan Dan Pengabdian Vokasi (JP2V)*, *1*(1), 44–57.

https://doi.org/10.32672/jp2v.v1i1.1840
Fahrurrozi, M., & Laili Rahmawati, S. N. (2021). Pengembangan model instrumen evaluasi menggunakan aplikasi kahoot pada pembelajaran ekonomi. *Jurnal PROFIT Kajian Pendidikan Ekonomi Dan Ilmu Ekonomi*, *8*(1), 1–10. https://doi.org/10.36706/jp.v8i1.13090
Faradillah, A., Hadi, W., & Soro, S. (2020). *Evaluasi proses dan hasil belajar (ephb) matematika dengan diskusi dan simulasi (disi) (1st ed.)*. Jakarta: Uhamka Press.
Fariyani, Q., Rusilowati, A., & Sugianto. (2015). Pengembangan four-tier diagnostic test untuk mengungkap miskonsepsi fisika siswa SMA kelas X. *Journal of Innovative Science Education*, *4*(2), 41–49. https://journal.unnes.ac.id/sju/index.php/jise/article/view/9903
Firdaus, D., & Rahayu, Y. S. (2020). Pengembangan instrumen tes miskonsepsi menggunakan four-tier multiple choice submateri katabolisme karbohidrat kelas XII SMA. *BioEdu*, *9*(2), 131–139. https://doi.org/10.26740/bioedu.v9n2.p131-139
Fitri, R., Syofyati, N., & Alberida, H. (2018). Undertanding's analysis the concept of classification of living organism for student's class vii at smpn 8 padang. *Bioeducation Journal*, *5*(2), 2021. https://doi.org/10.24036/bioedu.v5i2.142
Fitriani, N., & Rohaeti, E. E. (2020). Miskonsepsi siswa pada materi geometri di tingkat sekolah menengah pertama. *Teorema: Teori Dan Riset Matematika*, *5*(1), 9–16. https://doi.org/10.25157/teorema.v5i1.3267
Gultom, D. L. S. (2019). Miskonsepsi siswa pada materi klasifikasi makhluk hidup di kelas vii mts. Swasta al-washliyah sigambal tahun pembelajaran 2018/2019. *Iomas*, *1*(2), 39–42. https://jurnal.ulb.ac.id/index.php/berkala/index
Hasanti, M. A., & Zulyusri, Z. (2021). Meta-analisis miskonsepsi siswa mata pelajaran IPA materi biologi tingkat SMP. *PENDIPA Journal of Science Education*, *6*(1), 263–268. https://doi.org/10.33369/pendipa.6.1.263-268
Intan, N., & Masriyah, M. (2020). Pemberian scaffolding terhadap miskonsepsi siswa dalam menyelesaikan soal materi himpunan. *MATHEdunesa : Jurnal Ilmiah Pendidikan Matematika*, *9*(1), 221–230. https://doi.org/10.26740/mathedunesa.v9n1.p221-230
Muluki, A., & Bundu, P. (2020). Analisis kualitas butir tes semester ganjil mata pelajaran ipa kelas iv mi radhiatul adawiyah. *Jurnal Ilmiah Sekolah Dasar*, *4*(1), 86–96. https://doi.org/10.23887/jisd.v4i1.23335
Kaltakci-Gurel, D., Eryilmaz, A., & McDermott, L. C. (2017). Development and application of a four-tier test to assess pre-service physics teachers' misconceptions about geometrical optics. *Research in Science and Technological Education*, *35*(2), 238–260. https://doi.org/10.1080/02635143.2017.1310094
Kameswari, D. (2022). Pengetahuan mahasiswa biologi terhadap penggunaan terminologi bahasa latin. *Research and Development Journal Of Education*, *8*(1), 256–262. http://dx.doi.org/10.30998/rdje.v8i1.11187
Kurtuluş, M. A., & Tatar, N. (2021). An analysis of scientific articles on science misconceptions: A bibliometric research. *Elementary Education Online*, *20*(1), 192–207. https://doi.org/10.17051/ilkonline.2021.01.022
Laliyo, L. A. R., Hamdi, S., Pikoli, M., Abdullah, R., & Panigoro, C. (2021). Implementation of four-tier multiple-choice instruments based on the partial credit model in evaluating students' learning progress. *European Journal of Educational Research*, *10*(2), 825–840. https://doi.org/10.12973/EU-JER.10.2.825
Lestari, A., & Susantini, E. (2020). Pengembangan instrumen tes miskonsepsi menggunakan four-tier test pada materi transpor membran. *BioEdu*, *9*(3), 371–377. https://doi.org/10.26740/bioedu.v9n3.p371-377
Madukubah, F., & Taiyeb, M. (2018). Identifikasi miskonsepsi siswa pada konsep substansi genetik dengan menggunakan three tier test di kelas XII sekolah menengah atas. *Proceeding Biology Education Conference*, *15*(1), 236–242. https://jurnal.uns.ac.id/prosbi/article/view/31766
Magdalena, I., Fauziah, S. N., Faziah, S. N., & Nupus, F. S. (2021). Analisis validitas, reliabilitas, tingkat kesulitan dan daya beda butir soal ujian akhir semester tema 7 kelas iii sdn karet 1 sepatan. *BINTANG : Jurnal Pendidikan Dan Sains*, *3*(2), 198–214. https://ejournal.stitpn.ac.id/index.php/bintang
Mariyadi, M., & WA, I. R. (2023). Analisis miskonsepsi peserta didik kelas vi sekolah dasar pada

pembelajaran ipa materi gaya gravitasi. *Lentera Sains: Jurnal Pendidikan IPA*, *13*(1), 77–85. https://doi.org/10.24929/lensa.v13i1.225
Mubarak, S., Susilaningsih, E., & Cahyono, E. (2016). Pengembangan Tes Diagnostik Three Tier Multiple Choice Untuk Mengidentifikasi Miskonsepsi Peserta Didik Kelas Xi. *Journal of Innovative Science Education*, *5*(2), 101–110. https://journal.unnes.ac.id/sju/index.php/jise/article/view/14258
Muzakki, N. A., Diana, S., & Priyandoko, D. (2023). Misconception Analysis of Cell Material Using Four-Tier Multiple Choice Diagnostic Test. *Quagga: Jurnal Pendidikan Dan Biologi*, *15*(2), 120–130. https://doi.org/10.25134/quagga.v15i2.32
Noerinda. (2014). Penerapan pembelajaran ipa terpadu di SMP menjelang implementasi kurikulum 2013. *Lembaran Ilmu Kependidikan*, *43*(1), 25–31. https://doi.org/10.15294/lik.v43i1.3168
Nofiana, M., Sajidan, S., & Puguh, P. (2014). Pengembangan instrumen evaluasi two-tier multiple choice question untuk mengukur keterampilan berpikir tingkat tinggi pada materi kingdom plantae. *Jurnal Inkuiri*, *3*(2), 60–74. https://jurnal.uns.ac.id/inkuiri/article/view/9694
Noor, T. R. (2021). Strategi Solutif Kepala Sekolah Pada Pembelajaran Daring Selama Pandemi Covid 19 di SDN Sumput Sidoarjo. *Al-Fikrah: Jurnal Manajemen Pendidikan*, *9*(1), 20. https://doi.org/10.31958/jaf.v9i1.2658
Nurdiansyah, I., Hartati, S., & Paujiah, E. (2022). Analisis miskonsepsi siswa pada materi evolusi sub-bab adaptasi menggunakan metode three tier test. *Jurnal Literasi Pendidikan Dasar*, *3*(2), 29–39. https://jurnal.unikastpaulus.ac.id/index.php/jlpd/article/view/2059
Nurfajrianti, A. (2021). Efektivitas media berbasis material lokal dalam meningkatkan keterampilan klasifikasi siswa. *Report of Biological Education*, *2*(1), 28–37. https://doi.org/10.37150/rebion.v2i1
Nurhalimah, S., Hidayati, Y., Rosidi, I., Madura, U. T., & Soal, K. (2022). Hubungan antara validitas item dengan daya pembeda dan tingkat kesukaran soal pilihan ganda PAS. *Jurnal Natural Science Educational Research*, *4*(3), 249–257. https://doi.org/10.21107/nser.v4i3.8682
Nurhidayah, L., Riandi, R., & Solihat, R. (2020). Identifikasi miskonsepsi siswa SMA pada topik ekosistem. *Assimilation: Indonesian Journal of Biology Education*, *3*(1), 12–17. https://doi.org/10.17509/aijbe.v3i1.23303
Nurul Humaidi, M., Agus Triansyah, F., Sugianto, R., & Rachma Nisfi Laila, A. (2023). Development of a hots-leveled two tier multiple choice (ttmc) test to measure student misconceptions in islamic studies. *Assyfa Journal of Islamic Studies* , *1*(1), 31–40. https://doi.org/10.61650/ajis.v1i1.148
Pakpahan, T. R., Hernawati, D., & Ardiansyah, R. (2020). Analisis miskonsepsi peserta didik pada materi sistem saraf menggunakan four-tier diagnostic test. *BIOEDUSCIENCE: Jurnal Pendidikan Biologi Dan Sains*, *4*(1), 27–36. https://doi.org/10.29405/j.bes/4127-364844
Prianidya, A. P., Wisanti, & Setiawan, B. (2015). Analisis miskonsepsi siswa smp kelas VII pada mata pelajaran IPS melalui metode Certainty Of Response Index (Cri). *Pensa: Jurnal Pendidikan Sains*, *3*(3), 1–8. https://www.neliti.com/id/publications/251625/analisis-miskonsepsi-siswa-smp-kelas-vii-pada-mata-pelajaran-ipa-melalui-metode
Prihatiningsih, Zubaidah, S., & Kusairi, S. (2016). Kemampuan berpikir kritis siswa SMP pada materi klasifikasi makhluk hidup. In *Prosiding Seminar Nasional Pendidikan IPA Pascasarjana UM* (Vol. 1).
Puspitasari, Y., Reza, S. P. K., Bachtiar, Y., & Prayitno, B. A. (2019). Identifikasi miskonsepsi materi jaringan tumbuhan pada mahasiswa pendidikan biologi di salah satu universitas di surakarta. *BIOEDUKASI (Jurnal Pendidikan Biologi)*, *10*(2), 171. https://doi.org/10.24127/bioedukasi.v10i2.2494
Putri, R. E., & Subekti, H. (2021). Analisis miskonsepsi menggunakan metode four-tier certainty of response index: studi eksplorasi di smp negeri 60 surabaya. *Pensa E-Jurnal: PENDIDIKAN SAINS*, *9*(2), 220–226. https://ejournal.unesa.ac.id/index.php/pensa/article/view/38279/33954
Putri, R. S., Wigati, I., & Laksono, P. J. (2022). Faktor-faktor yang mempengaruhi miskonsepsi siswa pada materi asam dan basa. *Prosiding Seminar Nasional Pendidikan Kimia*, 280–286. https://proceedings.radenfatah.ac.id/index.php/snpk/article/view/80
Ramdani, A., Jufri, A. W., Gunawan, Fahrurrozi, M., & Yustiqvar, M. (2021). Analysis of students' critical thinking skills in terms of gender using science teaching materials based on the 5e learning cycle integrated with local wisdom. *Jurnal Pendidikan IPA Indonesia*, *10*(2), 187–199. https://doi.org/10.15294/jpii.v10i2.29956
Ramdani, A., Jufri, A. W., Gunawan, G., Hadisaputra, S., & Zulkifli, L. (2019). Pengembangan alat evaluasi

pembelajaran IPA yang mendukung keterampilan abad 21. *Jurnal Penelitian Pendidikan IPA*, *5*(1), 98–108. https://doi.org/10.29303/jppipa.v5i1.221
Sa'diyah, D. (2023). Analisis kebutuhan awal pengembangan bahan ajar ipa bagi siswa sekolah dasar. *Jurnal Bintang Pendidikan Indonesia (JUBPI)*, *1*(1), 11–17. https://doi.org/10.55606/jbpi.v1i1.917
Samiha, Y. T., Agusta, E., & Gestri, R. (2017). Analisis miskonsepsi siswa pada mata pelajaran IPA di SMPN 4 Penukal Utara Kabupaten Penukal Abab Lematang Ilir Pendopo. *Bioilmi*, *3*(1), 38–46. https://doi.org/10.19109/bioilmi.v3i1.5167
Sari, Y., Irawati, R. K., & Rahmawati, H. (2023). Practicality of web-based four-tier test to identify student's misconceptions in chemical bonding materials. *JPPS (Jurnal Penelitian Pendidikan Sains)*, *12*(2), 108–121. https://doi.org/10.26740/jpps.v12n2.p108-121
Soeharto, Csapó, B., Sarimanah, E., Dewi, F. I., & Sabri, T. (2019). A review of students' common misconceptions in science and their diagnostic assessment tools. *Jurnal Pendidikan IPA Indonesia*, *8*(2), 247–266. https://doi.org/10.15294/jpii.v8i2.18649
Sumardi. (2020). *Teknik pengukuran dan penilaian hasil belajar [Measurement techniques and assessment of learning outcomes].* Deepublish.
Suprapto, N. (2020). Do we experience misconceptions?: An ontological review of misconceptions in science. *Studies in Philosophy of Science and Education*. https://scie-journal.com/index.php/SiPoSE/article/view/24
Syaifuddin, M., Darmayanti, R., & Rizki, N. (2022). Development of a two-tier multiple-choice (ttmc) diagnostic test for geometry materials to identify misconceptions of middle school students. *Jurnal Silogisme*, *7*(2), 66–76. http://journal.umpo.ac.id/index.php/silogisme
Syarif, N., Alberida, H., Fitri, R., & Yogica, R. (2023). Identifikasi miskonsepsi materi sel pada peserta didik di kelas xi ipa man 2 kota padang. *Bioilmi : Jurnal Pendidikan*, *IX*(2), 45–52. https://doi.org/10.19109/bioilmi.v9i2.16270
Treagust, D. F. (1988). Development and use of diagnostic tests to evaluate students' misconceptions in science. *International Journal of Science Education, 10*(2), 159–169. https://doi.org/10.1080/0950069880100204
Widiastutik, E., & Isnawati. (2021). Profil miskonsepsi siswa kelas XII SMA pada submateri sintesis protein berdasarkan hasil uji four-tier diagnostic test. *BioEdu: Berkala Jurnal Pendidikan Biologi*, *10*(1), 85–94. https://doi.org/10.26740/bioedu.v10n1.p85-94
Wulandari, S., Gusmalini, A., & Zulfarina. (2021). Analisis miskonsepsi mahasiswa pada konsep genetika menggunakan instrumen four tier diagnostic test. *Jurnal Pendidikan Sains Indonesia*, *9*(4), 642–654. https://doi.org/10.24815/jpsi.v9i4.21153
Wulandari, S. I., Pamelasari, S. D., & Hardianti, R. D. (2023). Penggunaan e-modul berbasis etnosains materi zat dan perubahannya dalam usaha meningkatkan kemampuan berpikir kritis siswa SMP. *Prosiding Seminar Nasional IPA*, 103–113. https://proceeding.unnes.ac.id/snipa/article/view/2294
Yasthophi, A., & Ritongga, P. S. (2019). Pengembangan instrumen test diagnostik multiple choice four tier pada materi ikatan kimia. *Konfigurasi : Jurnal Pendidikan Kimia Dan Terapan*, *3*(1), 23. https://doi.org/10.24014/konfigurasi.v3i1.6797
Yusriya, A., Santosa, K., & Priyono, B. (2014). Pengembangan video pembelajaran materi klasifikasi hewan sebagai suplemen bahan ajar biologi SMP. *Journal of Biology Education*, *3*(1), 26–34. https://doi.org/10.15294/jbe.v3i1.4152