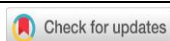


A Four-tier Test to Identify Students' Conceptions in Inheritance Concepts

Noviah Rosa Firdaus¹, Tjandra Kirana², Endang Susantini³
^{1,2,3} Universitas Negeri Surabaya, Surabaya, Indonesia



DOI : <https://doi.org/10.46245/ijorer.v2i4.128>

Sections Info

Article history:

Submitted: June 29, 2021

Final Revised: July 7, 2021

Accepted: July 10, 2021

Published Online: July 31, 2021

Keywords:

Four-tier Test

Inheritance

Students' Conceptions



ABSTRACT

Identification of students' conceptions is an important step before developing suitable learning methods to improve students' understanding. One of the best way to identify students' conception is the use of diagnostic test. Therefore, the development of diagnostic tests were important in the teaching and learning process to help the teacher determine the students' conceptions, wheater they have scientific conceptions, lack of knowledge, or misconception. The aim of this research is to develop a valid and reliable four-tier diagnostic test to identify students' conceptions in inheritance. The research method consists of three main stages: define content, obtaine information about students' conception, and develop the four-tier test. The result from the research demonstrated that the four-tier test that was developed, was valid and reliable instrument in diagnosing students' conception in inheritance concepts.

INTRODUCTION

Science is something related to activities that describe, predict, and find explanations for events that occur in the world. The results of science can be represented as an entity that is believed or can be analyzed, called a concept. Concepts that are built in students' minds are obtained from the learning process to make those learning meaningful (Sarwoto et al., 2020). But, everyone has a different understanding concept. Individual understanding of a concepts are referred to a conception. Sometimes, since there are many concepts in science, students write some answers that don't match to the correct concepts (Mimanah et al., 2020). Conceptions can be divided into several categories including understanding scientific conceptions, lack of knowledge, and misconceptions (Kaltakci-Gurel et al., 2017).

The process of identifying and determining the students' profile conceptions can be carried out during learning, because during learning process students are doing an understanding of the concept (Uliyandari et al., 2021)The process involves teacher-developed assesmennts or can be adapted from reliable sources. The diagnose of the students' conception profiles can't be done by using ordinary assesments technique but it can be done by using diagnostic assessment. Diagnostic assesment or usually called diagnostic test is an assessment tool that focus to persistent, recurring, and uresolved learning difficulties (Kaltakci-Gurel, 2015). A diagnostic assement tool is an spesific assessment that can be used to identify the competencies, strengthness, and weaknesses of students so that learning can be designed according to the competencies and conditions of students (Kepmendikbud, 2020). In Kepmendikbud (Decree of the Minister of Education and Culture Republic of Indonesia) Number 719 that was released in 2020 stated that the learning process must begin by giving a diagnostic test hence the teacher can develop the best learning methods or strategies.

Several forms of diagnostic tests that have been developed include: interviews (Jusniar et al., 2020; Duda & Apriyadi, 2020; Herman et al., 2011), concept maps (Plotz, 2020), drawing (Aksan et al., 2015), multiple-choice (Lamichhane et al., 2018), and multiple-choice with tier that are two-tier (Jung, 2020), three-tier (Jusniar et al., 2020), and four-tier (Kaniawati et al., 2019). Each form of the diagnostic test has advantages and disadvantages (Wola, 2020). Interview are the most important technique that can be used to determine students' conceptions to find out peoples' mind and what they think but it consumes a lot of time to be conducted (Kaltakci-Gurel et al., 2015). Multiple-choice tests can be used for large samples, easier to be analyzed, but can't investigate students' response in depth (Haladyna et al., 2019; Kaltakci-Gurel et al., 2015; Karaoglan-Yilmaz et al., 2020). The traditional multiple-choice test also can't distinguish the correct answer for the right reason or the wrong reason or the student might only guess, so it can't guarantee the correct answer because students have the correct understanding and follow scientific concepts (Türkoguz, 2020; Kaltakci-Gurel et al., 2015). On the other hand, wrong answer on traditional multiple-choice tests can't guarantee that misconceptions occur but it may be caused by students choose the wrong answers with the correct background.

The advantages of traditional multiple-choice tests make it necessary to develop these tests to overcome the shortcomings of diagnostic tests in the form of interviews, essay tests, and traditional multiple-choice tests. This is done by making multiple-choice tests with many tier, including two-tier, three-tier, and four-tier. As with other forms of diagnostic tests, the three stratified multiple-choice tests that have been developed by researchers have advantages and disadvantages. The two-tier multiple-choice test is a multiple-choice test that was first developed, then refined with a three-tier multiple-choice, and then a four-tier multiple-choice test was developed to overcome the shortcomings of the previous multiple-choice test. The four-tier multiple-choice test is a test that has been developed from the three-tier multiple-choice test which consists of four levels, first level was multiple-choice test, second tier was students' confidence in answering first tier, third tier was the reason for answering the first tier, and the fourth tier was students' confidence in choosing the reasons at the third level (Kaltakci-Gurel et al., 2015). The four-tier diagnostic test can correctly diagnose students' conception (Laliyo et al., 2020).

One of the branches of Biology is genetics which studies various ways of inheritance of traits to their offspring and develops into a more complex science and becomes the basis of other sciences, including biotechnology (Stern, 2017). Scientists consider genetics as the key to providing solutions to various disease that develop today's society (William et al., 2012). Furthermore, teachers or educators need to determine how students build relationships between complex ideas on genetics and the knowledge to understand the genetic lessons. Some researches about students' understanding of inheritance have been carried out by much previous research. It has shown that students have many difficulties while solving about inheritance concepts.

The study was conducted by Marbach-Ad et al (2000) reveal that students are usually unable to differentiate between genotype and phenotype. Karagoz et al (2011) stated that students experienced misconceptions on the basic concepts of inheritances, including differences in alleles and genes, dominance events and epistasis. The research that was conducted by William et al (2012) showed that students had difficulty connecting concepts in inheritance, especially using Punnet diagrams to predict and explain how homozygous offspring for the inheritance of certain traits could have two

heterozygous parents for the same trait. The achievement of the National Exams from 2016 until 2019 in genetics showed that students' genetic value percentage below the standard set by the Ministry of Education, which is only 45%, while the minimum standard is 55% (Puspendik, 2021). The researchers' findings through direct observation of twelfth grader students in various high schools in Surabaya showed that students have difficulty in solving problems in the inheritance of traits. Students consider the material to be complicated because there are many interrelated concepts.

The students' difficulties experience when studying the inheritance concepts make it increasingly important to carry out diagnostic tests for students. It can be revealed students' understanding of the concepts that exist in the inheritance concepts so that teachers can develop better learning methods. Therefore, the researcher conducted the research with the title "A Four-tier Test to Identify Students' Conceptions in Inheritance Concepts".

RESEARCH METHODS

The research is a research and development methods consist of two main stage. The first research stage was to develop a four-tier test referring to the research of Kilic and Saglam (2009) with several modifications. There were three main stages: defined the material, collected various information about students' conceptions, and developed the four-tier test. The four-tier test then validated based on content and criterion-related evidence. Content validation is carried out by three validators by giving a score of 1 if it is in accordance with the assessment indicators and a score of 0 if it is not. Assessment indicators on content validation were developed by researchers based on Tridianti (2014) and consist of three main aspects, namely material aspects, question construction, and language with a total 24 of indicators. Criterion validation is done by calculating the correlation coefficient (r) with the SPSS 25 version application. The four-tier test also measured about its reliability. The reliability of the four-tier test is measured by using the Cronbach alpha coefficient with the The reliability of the four-tier test was measured using the Cronbach alpha coefficient with the SPSS 25 version application. The second research stage was to administered the students' conception using the four-tier diagnostic test which done developed. The four-tier test administer to ten twelfth grader students of 4 Sidoarjo Senior High School.

Four-tier Test Development

1) Defining the Material. The material defined by semi-formal interviews with high school teachers in several schools in Sidoarjo and Suarabaya about which materials were the most difficult according to students. Researcher also analyzed the data on National Exam in 2016-2019 and conducted direct observations on approximately 60 twelfth grader students from several high schools in Surabaya. 2) Collection Various Information about Students' Conceptions. Researchers examine in-depth various published articles related to stundents' conceptions. The purpose of this stage is to collect any students argument to become a distractors in four-tier test. 3) Develop Four-tier Test. The preparation of the diagnostic test begins with making multiple choice questions with answer options containing three distractors and one correct answer. These distractors are statements that researchers have found when reviewing previous studies on students' conceptions of inheritance concepts. Then the researcher developed a reason consisting of four choice options consisting of three distractor statements and one correct statement.

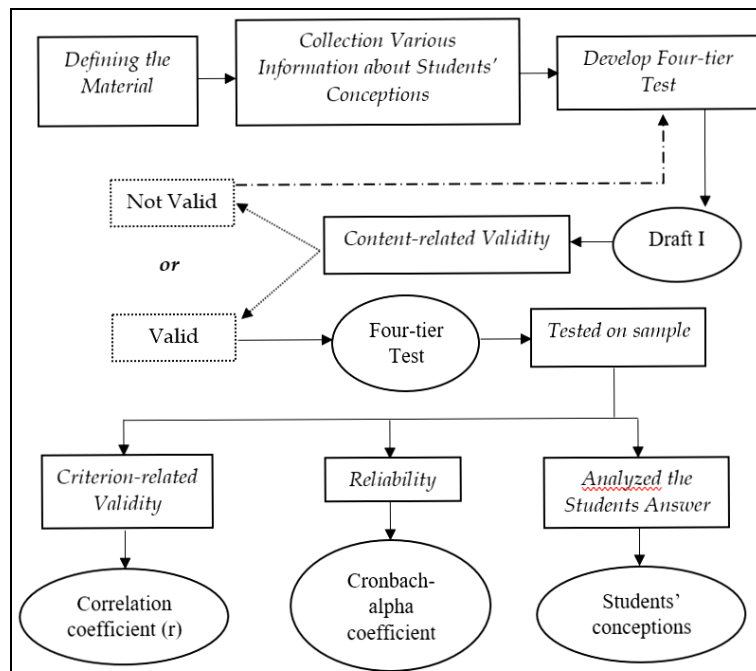


Figure 1. Research method.

Four-tier Test Validity

The purpose of validation is to assess the feasibility of the four-tier test. The four-tier test that has been developed then reviewed and validated by the three experts based on content-related evidence. The validation results calculated and analyzed according to the validity category. The following is the formula used to measure the validity of the four-tier test.

$$R = \frac{S}{N} \times 100\%$$

The value of validity then analyzed by validity category table:

Table 1. Validity score category.

Percentage (%)	Category
≤ 25 - 38	Highly invalid
39 - 25	Invalid
53 - 69	Quite valid
70 - 83	Valid
84 - 100	Highly Valid

Criteria validation is done by calculating the correlation coefficient (r) using SPSS 25 version. The results of the calculated r are compared with the r table. If r arithmetic is smaller than r table then the test is said to be invalid, whereas if r arithmetic is greater than r table then the test is said to be valid.

Four-tier Test Reliability

The four-tier test also measured for its reliability with Cronbach-alpha coefficient. The value of Cronbach-alpha coefficient then analyzed by reliability category table (table 2).

Table 2. Reliability category.

Reliability Coefficient	Category
$r_{11} \leq 0,20$	Highly unreliable
$0,20 \leq r_{11} < 0,40$	Unreliable
$0,40 \leq r_{11} < 0,60$	Quite reliable
$0,60 \leq r_{11} < 0,80$	Reliable
$0,80 \leq r_{11} < 1,00$	Highly Reliable

Students' Conceptions Profile

The four-tier test that was developed and validated then tested on ten twelfth grader students in 4 Sidoarjo Senior High School. The student's conception profile was determined from the score that was obtained on the four-tier test and then analyzed according to the conception category table. The following table was the category of students' conceptions with the four-tier test.

Table 3. Category of students' conceptions with the four-tier test.

Conceptions Category	Answer Combinations			
	First Tier	Second Tier	Third Tier	Fourth Tier
Scientific conception	True	Sure	True	Sure
	True	Sure	True	Not sure
	True	Not sure	True	Sure
	True	Not sure	True	Not sure
	True	Sure	False	Not sure
	True	Not sure	False	Sure
Lack of knowledge	True	Not sure	False	Not sure
	False	Sure	True	Sure
	False	Sure	True	Not sure
	False	Not sure	True	Sure
	False	Not sure	True	Not sure
	False	Sure	False	Not sure
Misconception	False	Not sure	False	Sure
	True	Sure	False	Sure
	False	Sure	False	Sure
	True	Sure	False	Sure

RESULTS AND DISCUSSION*Four-tier Inheritance Test*

This research has produced 12 questions of four-tier inheritance diagnostic test. Each question assesses a concept of inheritance. The first tier is multiple choice with four answer options that consist of one correct answer and three distractors. The second tier is students' belief in answering the test that has two options namely "sure" and "not sure". The third tier is the choice of reasons with four answer options which also consist of one correct reason and three distractors. The fourth tier is the student's belief in choosing a reason which also consists of two choices, namely "sure" and "not sure". Here is one of the questions that has been developed.

misinterpreted by students and an answer that is completely wrong. The existence of a level of confidence is one form of changing the structure of a multiple choice test that can determine the accuracy of respondents in answering questions, whether they are just guessing or really understanding the concepts in the questions.

Validity and Reliability of Four-tier Inheritance Test

The four tier test diagnostic test that has been developed is then reviewed and validated by three experts to assess the content-related validation and feasibility of the test. Content-related validation includes three aspects: material aspects, question construction, and language. In the material aspect, there are four things that are assessed. The results of the validation of the three experts on the material aspect are presented in Table 4.

Table 4. The percentage of validation on the aspects of the developed four-tier test diagnostic test material.

Number of Question	The total score each Experts			Percentage	Category
	1	2	3		
1	4	4	4	100%	Highly Valid
2	4	4	4	100%	Highly Valid
3	4	4	4	100%	Highly Valid
4	4	4	4	100%	Highly Valid
5	4	3	4	92%	Highly Valid
6	4	4	4	100%	Highly Valid
7	3	4	4	92%	Highly Valid
8	4	3	4	92%	Highly Valid
9	3	4	4	92%	Highly Valid
10	3	4	4	92%	Highly Valid
11	4	4	4	100%	Highly Valid
12	4	4	4	100%	Highly Valid
	Mean			96,7%	Highly Valid

Based on the Table 4, it can be observed that the feasibility of the material aspects of all the developed four-tier test diagnostic test items was declared highly valid with a feasibility average percentage of 96.7%. And every item have a percentage minimal 92%. It also means that the distribution is also evenly distributed on each item or number of question. The validation on the construction aspect of the question consists of 16 items that are assessed for feasibility. The following are the results of the construction validity of the three experts on each of the items developed.

Table 5. The percentage of validation on construction aspects of four-tier test diagnostic tests that have been developed.

Number of Question	Expert			Percentage	Category
	1	2	3		
1	14	15	14	90%	Highly Valid
2	13	15	14	88%	Highly Valid
3	11	15	14	83%	Valid
4	16	16	14	96%	Highly Valid
5	16	16	14	96%	Highly Valid
6	14	15	14	90%	Highly Valid
7	16	16	15	98%	Highly Valid

Number of Question	Expert			Percentage	Category
	1	2	3		
8	16	16	14	96%	Highly Valid
9	14	15	14	90%	Highly Valid
10	14	15	15	92%	Highly Valid
11	15	16	14	94%	Highly Valid
12	15	15	14	92%	Highly Valid
Mean				92,1%	Highly Valid

Based on the Table 5, it can be observed that the feasibility of the material aspects of all the developed four-tier test diagnostic test items was declared very valid with a feasibility percentage of 96.7%. Language validation consists of four things that are assessed for feasibility. The following are the results of the language validity of the three experts on each of the items that have been developed.

Table 6. The percentage of validation on construction aspects of four-tier test diagnostic tests that have been developed.

Number of Question	Experts			Percentage	Category
	1	2	3		
1	14	15	14	90%	Highly Valid
2	13	15	14	88%	Highly Valid
3	11	15	14	83%	Valid
4	16	16	14	96%	Highly Valid
5	16	16	14	96%	Highly Valid
6	14	15	14	90%	Highly Valid
7	16	16	15	98%	Highly Valid
8	16	16	14	96%	Highly Valid
9	14	15	14	90%	Highly Valid
10	14	15	15	92%	Highly Valid
11	15	16	14	94%	Highly Valid
12	15	15	14	92%	Highly Valid
Mean				92,1%	Highly Valid

Based on the Table 6, it can be seen that the language aspect of the 11 four-tier diagnostic test items that has been developed was declared highly valid and one item was declared valid. The feasibility of the four-tier diagnostic test on the language aspect was declared highly valid with a 96% eligibility percentage. The results of the validation by three experts on the three aspects assessed were declared highly valid with a feasibility percentage of 93%. Criterion-related validation dilakukan dengan menghitung koefisien korelasi (r) dengan SPSS versi 25. The criterion-related validation calculation results are described in Table 7.

Table 7. Correlation coefficient.

Question	r^*	r_{table}	Category
1	0,663	0,632	Valid
2	0,875	0,632	Valid
3	0,663	0,632	Valid
4	0,663	0,632	Valid
5	0,875	0,632	Valid
6	0,663	0,632	Valid
7	0,855	0,632	Valid

Question	r *	r _{table}	Category
8	0,875	0,632	Valid
9	0,875	0,632	Valid
10	0,677	0,632	Valid
11	0,875	0,632	Valid
12	0,663	0,632	Valid
Mean	0,7685	0,632	Valid

*Correlation is significant at the 0,01 level

The results of criterion-related validation showed that the twelve questions which were developed and tested to students are valid questions. This showed that the four-tier test of inheritance material in high school can be used as an instrument to define students' conceptions of inheritance material. In the other words, that the data obtained from the test can validly be used as a conclusion. The reliability of the four-tier test was calculated using the Cronbach-alpha approach. The following is a recapitulation of the reliability of the four-tier test that has been developed.

Reliability Statistics	
Cronbach's Alpha	N of Items
.942	12

Figure 3. Cronbach-alpha coefficient.

Based on the recapitulation, it can be seen that the reliability coefficient calculated by the Cronbach-alpha approach is 0.942 and is categorized reliable. This means that the questions that has been developed are quite consistent if they are done by the respondents.

Students' Conceptions about Inheritance Concepts

The four-tier test that has been developed can be an instrument to identify the student's conception profile on the inheritance of traits. The student's overall conception of the inheritance of traits is presented in the Figure 4.

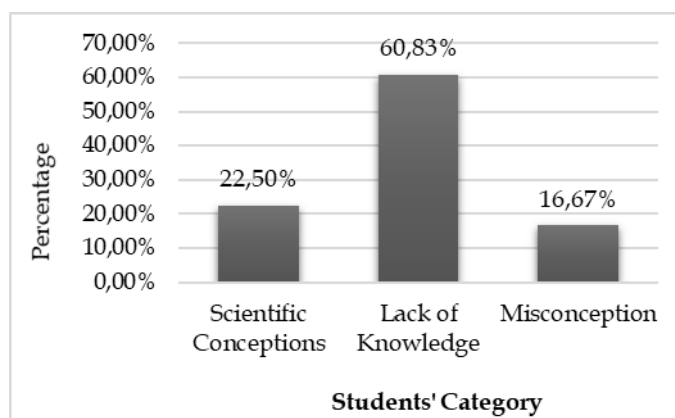


Figure 4. Students' conceptions.

Each student builds understanding in a different way and the process to get that understanding might be varies depending on the conceptual understanding that

students have (Aktan, 2013). Those makes each student have a different level of learning progress, level of understanding of a concept, and frame of mind (Laliyo et al., 2021). Students who have difficulty understanding on certain concepts will make them feel difficult to build explanations related to basic science concepts. Using diagnostic tests will provide reliable information related to the condition of students at that time thus it help teachers develop accurate and valid learning components to guide students to improve their abilities.

The students' conceptions in this study were divided into three categories, namely scientific concepts, lack of knowledge and misconceptions. Based on Figure 4 it is known that the highest student conception is not understanding the concept. More than half of the students that were tested did not understand the concepts in the inheritance of traits. The lowest percentage of conceptions is misconceptions. Misconceptions only occurred in some students. The percentage of students who understand the concept is higher than students who experience misconceptions but it is still far below the percentage of students who do not understand the concept. The number of students who understand the concept is more than students who experience misconceptions.

The conception category "Understanding" the concept is defined from the translation of scientific conception or scientific conception which means students' understanding of a concept that is in accordance to what has been agreed upon by scientists (Kampourakis, 2018). Students can be concluded as understanding the concept or having a scientific conception, then the answers given are followed by various attributes of the correct concept. Each concept has five important attributes, namely name, definition, specific characteristics, examples, and values (Ibrahim, 2012). Based on the concept category table (table 3) students can be concluded as understanding the concept if the answer is correct at the first stage, answers "sure" at the second stage, the answer is correct at the third stage, and answers "sure" at the fourth stage.

Not understanding the concept means that the student has got the concepts but he has not understood the meaning. Student's understanding to a concept is a student's thinking process to proceed what they have received when learning so that it becomes meaningful (Fatimah, 2017). Students who can process various information in the learning process can understand the concept correctly. Students who have understood a concept will show better self-confidence so that they can choose "sure" at the second and fourth levels. The existence of this understanding makes students easily to solve or answer the questions that are given.

Another category of student conceptions is student misconceptions that occur in inheritance (16.67%). Research that is related to misconceptions on the concepts of inheritance has also been widely carried out and has determined the existence of misconceptions in various concepts, including patterns of inheritance (Osman et al., 2017), the use of Punnet diagrams in crosses (William et al., 2012).), and terms of inheritance (Stern et al., 2020; Stern et al., 2017). Most students said that the concepts of inheritance contain a lot of scientific terms and students tend to memorize these terms rather than understand them.

Misconception has many synonyms which mean the same thing, namely misunderstanding. Many literatures refer to misconceptions as alternative conceptions, nave conceptions, and misconceptions (Osman et al., 2017). Misconception is defined as a conceptual pattern within the individual that deviates from the meaning widely accepted by the scientific community or scientist . Misconceptions are usually resistant and do not change using traditional teaching methods. Misconceptions that occur in

students can be happened because of students' views or perspectives on a concept. In addition, it can also occur because of an experience they have found in everyday life. Formal and non-formal education can also be a source of students experiencing misunderstandings. Formal education includes teachers or tutors and learning resources used (Suwono et al., 2019). Yates et al (2014) also showed that teachers can be a source of misconceptions for students.

CONCLUSIONS

The development of diagnostic tests is very important to diagnose students' understanding of a concept in depth. Student understanding is a very important component in a learning process. Teachers who already know student's understanding can develop appropriate learning methods and models so that learning objectives are achieved and misconceptions that are experienced by students can be corrected. In this study, a four-tier diagnostic test was produced to identify the student's conception profile on the inheritance material for grade twelfth grader senior high school students which were declared valid and quite reliable. The result of four-tier test can diagnose students' conceptions and distinguish them into understanding concepts, not understanding concepts, and misconceptions. The profile of students' conceptions in high school inheritance material consists of three categories of conception, namely understanding the concept; do not understand the concept; and misconceptions. This research only focused on the development the four-tier diagnostic test and not explained what are the conceptions on students' mind. For further research, it can be focused on developing learning models that are in accordance with the results obtained in this study.

REFERENCES

- Aksan, Z., & Celikler, D., (2015). Evaluation of the knowledge and misconceptions of science teacher candidates in turkey regarding the greenhouse effect through the use of drawings. *Journal of Education and Practice*, 6(13), 112-120.
- Aktan, D. C. (2013). Investigation of students' intermediate conceptual understanding levels: The case of direct current electricity concepts. *European Journal of Physics*, 34(1), 33-43. doi.org/10.1088/0143-0807/34/1/33
- 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. doi.org/10.1007/s11165-009-9122-4
- Duda, H. J., & Adpriaydi. (2020). Students' misconception in concept of biology cel. *Anatolian Journal of Education*, 5(1), 47-52. doi.org/10.29333/aje.2020.515a
- Fatimah, A. S., & Santiana, S. (2017). Teaching in 21st century: Students-teachers' perceptions of technology use in the classroom. *Script Journal: Journal of Linguistic and English Teaching*, 2(2), 125-135. doi.org/10.24903/sj.v2i2.132
- Habiddin, & Page, E. M. (2019). Development and validation of a four-tier diagnostic instrument for chemical kinetics (FTDICK). *Indonesian Journal of Chemistry*, 19(3), 720-736. doi.org/10.22146/ijc.39218
- Haladyna, T. M., Rodriguez, M. C., & Stevens, C. (2019). Are multiple-choice items too fat?. *Applied Measurement in Education*, 32(4), 350-364. doi.org/10.1080/08957347.2019.1660348

- Herrmann-Abell, C. F., & Deboer, G. E. (2016). Using rasch modeling and option probability curves to diagnose students' misconceptions. *American Educational Research Association*, 8(12), 1-12. doi.org/10.1039/C6RP00004E
- Hoe, K. Y., & Subramaniam, R. (2016). On the prevalence of alternative conceptions on acid-base chemistry among secondary students: insights from cognitive and confidence measures. *Chemistry Education Research and Practice*, 17(2), 263-282. doi.org/10.1039/c5rp00146c
- Ibrahim, M. (2012). *Konsep, miskonsepsi, dan cara pembelajarannya*. Surabaya: Unesa University Press.
- Jung, J. (2020). Diagnosing causes of pre-service literature teachers' misconceptions on the narrator and focalizer using a two-tier test. *Education Science*, 10(4), 104-124. doi.org/10.3390/educsci10040104
- Jusniar, J., Effendy, E., Budiasih, E., & Sutrisno, S. (2020). Misconceptions in rate of reaction and their impact on misconceptions in chemical equilibrium. *European Journal of Educational Research*, 9(4), 1405-1423. doi.org/10.12973/eujer.9.4.1405
- Kaltakci-Gurel, D., Eryilmaz, A., & McDermott, L. C. (2015). A review and comparison of diagnostic instruments to identify students' misconceptions in science. *Eurasia Journal of Mathematics, Science & Technology Education*, 11(5), 989-1008. doi.org/10.12973/eurasia.2015.1369a.
- 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 & Technological educaTion*, 35(2), 238-260. doi.org/10.1080/02635143.2017.1310094
- Kampouraskis, K. (2018). On the meaning of concepts in science education. *Science & Education*, 27(1), 591-592. doi.org/10.1007/s11191-018-0004-x
- Karaoglan-Yilmaz., Ustun, A. B., & Yilmaz, R. (2020). Investigation of pre-service teachers' opinions on advantages and disadvantages of online formative assessment: An example of online multiple-choice exam. *Journal of Teacher Education and Lifelong Learning*, 2(1), 1-8.
- Kemendikbud. (2020). *Keputusan menteri pendidikan dan kebudayaan republik indonesia nomor 719/p/2020 tentang pedoman pelaksanaan kurikulum pada satuan pendidikan dalam kondisi khusus*. Kementerian Pendidikan dan Kebudayaan Republik Indonesia
- Kilic, D., & Saglam, N. (2009). Development of a two-tier diagnostic test concerning genetics concepts: the study of validity and reliability. *Procedia - Social and Behavioral Sciences*, 1(1), 2685-2686. doi.org/10.1016/j.sbspro.2009.01.474
- Laliyo, L. A. R., Hamdi, S., Pikoli, R., Abdullah, M., & 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. doi.org/10.12973/eu-jer.10.2.82
- Lamichhane, R., Reck, C., & Maltese, A., V. (2018). Undergraduate chemistry students' misconceptions about reaction coordinate diagrams. *Chem. Educ. Res. Pract.*, 19(1), 834-845. doi.org/10.1039/C8RP00045J
- Marbach-Ad, G., & Stavy, R. (2000). Students' cellular and molecular explanations of genetic phenomena. *Journal of Biological Education*, 34(4), 200-205. doi.org/10.1080/00219266.2000.9655718
- Mimanah, I. I. A., Suryanti, & Suprpto, N. (2020). Development of an inquiry based science learning material using flash card to reduce misconception of

- elementary school students. *IJORER*, 1(2), 178-190. doi.org/10.46245/ijorer.v1i2.35
- Osman, E., BouJadoude, S., & Hamdan, H. (2017). An investigation of lebanese G7-12 students' misconceptions and difficulties in genetics and their genetics literacy. *Int J of Sci and Math Educ*, 15, 1257-1280. doi.org/10.1007/s10763-016-9743-9
- Plotz, T. (2020). Are concept maps a valid measurement tool for conceptual learning? A cross-case study. *EURASIA Journal of Mathematics, Science and Technology Education*, 16(1), 1795-1816. doi.org/10.29333/ejmste/110174
- Puspendik. (2020). *Laporan hasil ujian nasional*. Jakarta: Pusat Penelitian Pendidikan
- Sarwoto, T. A., Jatmiko, B., & Sudiby, E. (2020). Development of online science teaching instrument based on scientific approach using phet simulation to improve learning outcomes at elementary school. *IJORER*, 1(2), 90-107. doi.org/10.46245/ijorer.v1i2.40
- Stern, F., Kampaourakis, K., Delaval, M. & Müller, A. (2020) Development and validation of a questionnaire measuring secondary students' genetic essentialism and teleology (GET) conceptions. *International Journal of Science Education*, 40(2), 218-252. doi.org/10.1080/09500693.2019.1707905
- Stern, F., & Kampaourakis, K. (2017). Teaching for genetics literacy in the post-genomic era. *Studies In Science Education*, 53(2), 193-225. doi.org/10.1080/03057267.2017.1392731
- Suwono, H et al. (2019). Cell Biology Diagnostic Test (CBD-Test) Portrays Pre-Service Teacher Misconceptions About Biology Cell. *Journal of Biological Education*, 55(2), 1-24. doi.org/10.1080/00219266.2019.1643765
- Türkoguz, S. (2020). Investigation of three-tier diagnostic and multiple choice tests on chemistry concepts with response change behaviour. *International Education Studies*, 13(9), 10-22. doi.org/10.5539/ies.v13n9p10
- Uliyandari, M., Candrawati, E., Herawati, A. A., & Latipah, N. (2021). Problem-based learning to improve concept understanding and critical thinking ability of science education undergraduate students. *IJORER*, 2(1), 65-72. doi.org/10.46245/ijorer.v2i1.56
- Williams, M., DeBarger, A. H., Montgomery, B. L., Zhou, X., & Tate, E. (2012). Exploring middle school students' conceptions of the relationship between genetic inheritance and cell division. *Science Education*, 96(1), 78-103.
- Wola, B. R., Ibrahim, M., & Purnomo, T. (2020). Development of a four-tier multiple-choice test on the concept of transport across membranes. *SEJ (Science Education Journal)*, 4(2), 77-97. doi.org/10.21070/sej.v4i2.878
- Yates, T. B., & Marek, E. A. (2014). Teachers teaching misconceptions: a study of factors contributing to high school biology students' acquisition of biological evolution-related misconceptions. *Evolution: Education and Outreach*, 7(7), 1-18. doi.org/10.1186/s12052-014-0007-2
- Zhao, Z. B. (2013). Diagnosing the english speaking ability of college students in china - validation of the diagnostic college english speaking test. *RELC Journal*, 44(3), 341-359. doi.org/10.1177/0033688213500581

***Noviah Rosa Firdaus, M.Pd. (Corresponding Author)**

Graduate Programs, Science Education,
State University of Surabaya,
Jl. Lidah Wetan, Surabaya, East Java, Indonesia
Email: noviah.19005@mhs.unesa.ac.id

Prof. Dr. dr. Tjandra Kirana, Sp.And.

Graduate Programs, Science Education,
State University of Surabaya,
Jl. Lidah Wetan, Surabaya, East Java, Indonesia
Email: nana.snoer@gmail.com

Prof. Dr. Endang Susantini, M.Pd.

Graduate Programs, Science Education,
State University of Surabaya,
Jl. Lidah Wetan, Surabaya, East Java, Indonesia
Email: endangsusantini@unesa.ac.id
