JPPIPA 8(4) (2022)

Jurnal Penelitian Pendidikan IPA

Journal of Research in Science Education

http://jppipa.unram.ac.id/index.php/jppipa/index

Development of Biological Macromolecules Three-Tier Test (BM-3T) to Identify Misconceptions of Prospective Science Teachers

Ni Wayan Suriani[1*], Brian Ricard Wola[1], Alfrits Komansilan[2]

[1] Department of Science Education, Universitas Negeri Manado, Minahasa, Indonesia.
[2] Department of Physics, Universitas Negeri Manado, Minahasa, Indonesia.

Article Info
Received: January 6, 2022
Revised: September 10, 2022
Accepted: October 20, 2022
Published: October 31, 2022

Corresponding Author:
Ni Wayan Suriani
niwayansuriani@unima.ac.id

DOI: 10.29303/jppipa.v8i4.1297

Abstract: This study aims to develop a valid and reliable biological macromolecules three-tier test (BM-3T) so that it is feasible to identify misconceptions of prospective science teachers. This research is a type of research and development by following the research design by Kiliç and Sağlam, which consists of three stages, namely defining content, obtaining information on student misconceptions, and developing tests. Research data was obtained through interviews, validation, and test trials. The subjects in this study consisted of 14 prospective science teachers in the first semester in the 2021/2022 academic year at the Universitas Negeri Manado. The data analysis carried out includes the analysis of test validity, test reliability analysis, and test difficulty level analysis. The results showed that the 16 test items developed were declared valid, the test reliability coefficient of 0.78 was in the reliable category, and the level of difficulty of BM-3T shows an average value of 0.44, which is included in the moderate category. Therefore, BM-3T is appropriate to identify misconceptions of prospective science teachers on the topic of biological macromolecules.

Keywords: Three-tier test; Biological macromolecules; Misconceptions; Prospective science teachers

Introduction

The General Biology I course is one of the subjects included in the science education study program at the Universitas Negeri Manado. The subject matter in General Biology I course provides knowledge of basic concepts and theories in the biological sciences (Katalog Jurusan Pendidikan IPA, 2021). One of the topics that are the subject of lecture studies in the General Biology I course is the topic of biological macromolecules. Living things that exist from the smallest body size to the enormous body size are composed of four macromolecules: carbohydrates, proteins, lipids, and nucleic acids. Biological macromolecules are polymers formed by linking monomers together through dehydration reactions. This process releases a water molecule for each bond formed (Mason et al., 2016).

The topic of biological macromolecules is a fundamental concept that must be studied by first-year science education undergraduate students because it is a piece of basic knowledge in understanding other topics. For example, the four classes of macromolecules are structural and functional components of cells. Therefore, the discussion of the structure related to its function is the material's content in the topic of biological macromolecules that most had discussed. Understanding the structure of biological molecules can provide important information about their functions and mechanisms of action (Jaswal et al., 2013). It is not surprising that a deep understanding of the structure of biological macromolecules forms the basis of research in various fields of science, including biotechnology, drug discovery, and disease therapy design (Boodhun, 2018).

Science is an inseparable part of our daily lives (Juliani et al., 2021). Studying science is one way for humans to understand natural phenomena that occur in life. Until now, misunderstandings or misconceptions about scientific concepts still occur (Gurel et al., 2015; Soeharto et al., 2019; Kurtulus & Tatar, 2021). Misconceptions about scientific concepts indicate that students' ability to understand concepts is still relatively low. Two things cause the low ability of students to understand concepts; namely, students do not understand the concept or misunderstand a concept, namely misconception (Adityawardani & Hidayati, 2017). Problems like this are not only experienced by

How to Cite:
Suriani, N.W., Wola, B.R., & Komansilan, A. (2022). Development of Biological Macromolecules Three-Tier Test (BM-3T) to Identify Misconceptions of Prospective Science Teachers. *Jurnal Penelitian Pendidikan IPA*, *8*(4), 1798-1805. https://doi.org/10.29303/jppipa.v8i4.1297

high school students but also experienced by students at the university level (Widiarti et al., 2019). Students usually have their understanding obtained in high school before studying in college. Therefore, knowledge is needed to identify and analyze how students understand concepts.

Misconceptions also occur in teachers (Arslan et al., 2012). Many pieces of research on the misconceptions of science teacher candidates have been carried out. The misconceptions of prospective science teachers in the field of physics include the topic of electricity (Zuhdi & Busyairi, 2021), force and motion (Al-Rsa'I, Khoshman, & Abu Tayeh, 2020), astronomy (Kanli, 2014; Ozkan & Ackay, 2016), and light (Wahyuni et al., 2019). The misconceptions of prospective science teachers in the field of chemistry include the topic of chemical bonds (Fatokun, 2016), chemical kinetics (Çam, Topçu, & Sülün, 2015), acids and bases (Safo-Adu, 2020), and intermolecular forces (Widiarti et al., 2019). In the field of biology, prospective science teachers experience misconceptions on the topics of plant classification (Yangin et al., 2014), evolution (Karataş, 2020), and photosynthesis (Karakaya et al., 2021), and respiration (Susanti, 2018). Nevertheless, research on the misconceptions of science teacher candidates on biological macromolecules topic is still rarely reported. Therefore, prospective teachers studying at university must first identify their conception profile to know about their misconceptions.

Students' conceptional profiles can be divided into three categories: understanding concepts, not understanding concepts, and misconceptions (Wola et al., 2020). Students are categorized as understanding concepts if the answers follow the approved scientific concepts. On the other hand, students are categorized as not understanding the concept if they give the wrong answer and are not sure or only guess the answer. Students who give wrong answers and certain the answers are categorized as misconceptions. Misconceptions are also referred to as alternative conceptions, naive beliefs, preconceptions, alternative frameworks, erroneous ideas, multiple private versions of science, personal models of reality, spontaneous reasoning, spontaneous knowledge, common-sense concepts, underlying sources of error, and children science (Yangin et al., 2014).

According to Suwarto (2013), a diagnostic test is needed to diagnose students' conceptions to find out students' conceptual understanding. Diagnostic tests are used to identify the profile of students' conceptions to distinguish students who understand the concept, do not understand the concept, and have misconceptions. Research conducted by Soeharto et al. (2019) reporting on diagnostic tests often used to identify student misconceptions in science learning in 111 scientific articles published from 2015 to 2019. Furthermore, the study showed that 10.74% used interviews, 23.97% used open-ended questions, 32.23% used multiple-choice, 9.92% used a *two-tier* multiple-choice *test*, 16, 53% used a *three-tier* multiple-choice test, 4.13% used a *four-tier* multiple-choice test, and 2.48% used another tiered test. This study indicates that the *three-tier* multiple-choice *test* is one of the diagnostic tests used to identify students' misconceptions.

The *three-tier* multiple-choice test is a diagnostic test consisting of three levels (Pesman & Eryilmaz, 2010). The first level is multiple choice, the second level is a choice of reasons for the first level, and the third level is a question of the level of confidence in the first and second levels (Yang & Sianturi, 2019; Liampa et al., 2019; Türkogus, 2020). This diagnostic test has students' belief items that provide more precise information about students' misconceptions. Also, it can distinguish between students who understand the concept, students who do not understand the concept or do not know the concept, and students who experience misconceptions (Gurel et al., 2015). In addition, the three-level multiple-choice test is also more valid in determining students' misconceptions than the standard multiple-choice test and the two-level multiple-choice test.

Based on interviews with lecturers for General Biology 1, it is known that the topic of biological macromolecules is complex for students to understand. The content of the lecture material is abstract, and many new terms are unfamiliar to first-year students, making it difficult for students to understand concepts. It is known from the results of student quizzes in previous years who scored below 3.00. The lecturers also never used any form of a diagnostic test, so it is clear that there has never been the identification of student misconceptions on the topics in this course. In this study, we aimed to develop a valid and reliable diagnostic test, namely the Biological Macromolecules Three-Tier Test (BM-3T), to identify misconceptions of prospective science teachers' on biological macromolecules. The test instrument we developed has the value of novelty due to the lack of reports in the extant literature relating to the usage of diagnostic tests to identify prospective science teacher misconceptions about biological macromolecules.

Method

This research is a type of research and development (R&D) by following the research design by Kiliç and Sağlam (2009), which consists of three stages, namely (1) defining content, (2) obtaining information on student misconceptions, and (3) developing tests as shown in Figure 1. The stage of defining content consists of field studies, determining concepts, taking an inventory of concepts, and reviewing the literature. Field studies were carried out through interviews with lecturers of the General Biology I course to decide what

topics would be identified and explore the existence of misconceptions. Next, the researcher takes an inventory of the concepts included in the essential concepts that have been determined previously in the form of a concept map. After that, the researcher conducted a literature review to describe the concepts inventoried to be used as references and collect information on misconceptions' characteristics. Decision-making about the category of misconceptions of science teacher candidates follows the type of conception profile by Arslan et al. (2012).

The stage of obtaining information about students' misconceptions is carried out by examining the misconceptions that have been previously reported in research articles and other scientific papers. In addition, the researchers also reviewed the test items on the quizzes that the lecturers used regarding biological macromolecules topics in previous years. Finally, in developing the test, the researchers compiled a test grid, arranged a *three-tier test* on biological macromolecules, and set the assessment criteria range to produce a draft I of the BM-3T. This draft was validated by three expert lecturers using a validation sheet to obtain a test validity score in terms of material aspects, construction aspects, and language aspects (Depdiknas, 2008), as shown in Table 1.

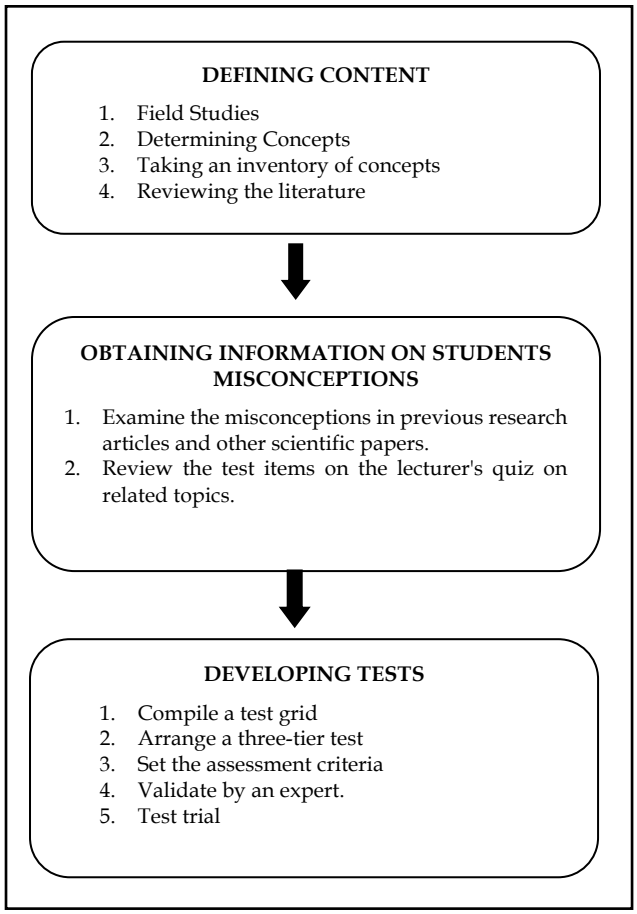

Figure 1. Stages of Research

Table 1. Test Item Validity Assessment Criteria

Aspect	Assessment criteria
Theory	1. Test items according to indicators
	2. Homogeneous and logical answer choices
	3. Choice of homogeneous and logical reasons
	4. Contents of material according to level, type of school, and grade level
Construction	5. The main questions are formulated briefly, clearly, and firmly
	6. Answer choices are formulated briefly and clearly
	7. The formulation of the main questions and the answer choices are required statements.
	8. The choice of reasons is formulated briefly and clearly
	9. The main question does not contain double negative statements
	10. The length of the answer choices is relatively the same
	11. The length of the choice of reasons is relatively the same
	12. Pairs of answer choices and reasons indicate an alternative to the occurrence of misconceptions
	13. Pictures, charts, the like are presented clearly
Language	14. Sentences using language that follows the rules of the Indonesian language
	15. Using communicative language, so easy to understand
	16. Do not use the local language
	17. The answer choices do not repeat the words or phrases contained in the main question.

Research data was obtained through interviews, validation, and test trials. The data analysis carried out includes the analysis of test validity, test reliability analysis, and test difficulty level analysis. Semi-structured interviews were conducted with the General Biology I course lecturers to get information about the subject matter potentially for misconceptions. The validity of the test is assessed based on the material, construction, and language aspects of each test item by three expert lecturers as validators. The validator responded by following a Likert scale of 1-4, namely 1 for disagreeing, 2 for opposing, 3 for agreeing, and 4 for strongly agreeing. Furthermore, the validity of the test items is known by calculating the average score of each test item so that we can see whether it is very valid, valid, less valid, or invalid. The validity of the test items is carried out based on the score interpretation criteria as shown in Table 2 (Riduwan, 2012). In addition, the validator also responds to each test item as revision material. Based on the responses written by the validator on the validation sheet, the test instrument was revised to produce a draft II BM-3T.

Table 2. Score Interpretation Criteria

Score	Category
3.51 - 4.00	Very Valid
2.51 - 3.50	Valid
1.51 - 2.50	Less Valid
1.00 - 1.50	Invalid

Based on these criteria, the developed test items are declared valid and suitable for use if they reach a minimum of 2.51. After all, the test items are scored, and the categories are determined, then determine the mode of the category of the validity of the test instrument as a whole, namely the number of test items that are very valid, valid, less valid, and invalid.

Draft II of BM-3T was piloted to 14 prospective science teachers in the first semester in the 2021/2022 academic year at the Universitas Negeri Manado after they studied the topic of biological macromolecules. The trial activity of the BM-3T draft II aims to obtain quantitative data in the form of the test reliability coefficient value and the level of difficulty of the test items. We use the KR-20 equation to determine the reliability coefficient, while the item difficulty level used the item difficulty equation suggested by Gronlund (1993). The KR-20 equation is as follows.

$$R_i = \frac{k}{k-1}\left\{\frac{S_t^2 - \sum p_i q_i}{S_t^2}\right\} \quad (1)$$

Description of equation (1), i.e., r_i is the coefficient of reliability of the instrument, k is the number of items in the instrument, p_i is the number of subjects who answered correctly, q_i is 1 - p_i, and S_t^2 is the total variance. Based on equation (1), it is clear that the total variance must be known in advance to calculate the reliability coefficient. The total variance correlation formula is as follows:

$$S_t^2 = \frac{\sum X_t^2 - \frac{(\sum X_t)^2}{N}}{N} \quad (2)$$

The description of equation (2), i.e., S_t^2 is the total variance, $\sum X_t$ is the total score, and N is the number of respondents.

The known value of the instrument reliability coefficient is then interpreted into the criteria, as shown in Table 3 (Arikunto, 2014). The developed BM-3T is reliable if it reaches a minimum reliability coefficient of 0.60.

Table 3. Interpretation Criteria for the Degree of Test Reliability

Reliability Coefficient	Interpretation
0.80 r_{11} 1.00	Very Reliable
0.60 r_{11} 0.80	Reliable
0.40 r_{11} 0.60	Reliable Enough
0.20 r_{11} 0.40	Not Reliable
R_{11} <0.20	Very Unreliable

The difficulty of the test items refers to the percentage of examinees who answered the test correctly (Zaman et al., 2010). The level of difficulty of the test items ranged from 0% to 100%. The level of difficulty of the test items uses the following equation (Gronlund, 1993).

$$P = \frac{R}{T} \times 100 \quad (3)$$

Information for equation (3), namely P is the percentage of test-takers who answered the test items correctly, R is the number of test-takers who answered the test items correctly, and T is the number of test-takers. The interpretation of the level of difficulty of the test items follows the categories in Table 4 (Karim, Sudiro, & Sakinah, 2021).

Table 4. Category of Test Item Difficulty Level

Index Range	Category
0.00 – 0.30	Difficult
0.31 – 0.70	Moderate
0.71 – 1.00	Easy

Result and Discussion

Using the three-tier test is highly recommended to diagnose misconceptions experienced by students (Jusniar et al., 2020). In this study, a diagnostic test in the form of a three-tier test that we developed aims to identify misconceptions of prospective science teachers on biological macromolecules topic. The first-tier contains multiple-choice with three answer choices. In the second tier, prospective science teachers can choose four reasons to confirm the answers given in the first tier. We asked prospective science teachers' beliefs about the first and second tiers' answers in the third tier. The level of confidence consists of two choices, namely sure and not sure. Sixteen test items in the BM-3T refer to the evaluation of concept elements, namely the concept name, definition, attributes (specific characteristics/essential characteristics), examples, and values. The BM-3T was developed in Indonesian, in which the test takers were most proficient. Figure 2 shows an example of test items.

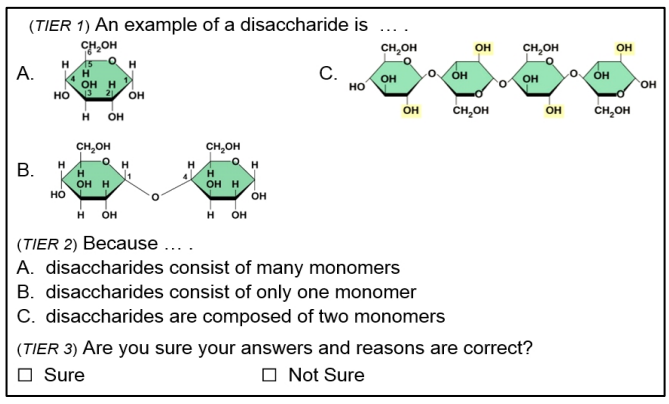

Figure 2. Example of Test Items on The BM-3T

A test is declared valid if the test can measure what should be measured (Arikunto, 2015). Three lecturers validated the BM-3T instrument from three different study programs at the Universitas Negeri Manado, experts in diagnostic tests, biology, and chemistry. The validator gives an assessment based on predetermined criteria, namely a value of 1 means disagrees, 2 means disagree, 3 means agree, and 4 means strongly agree. The validation results provide information about the level of validity as a reference for the feasibility of BM-3T in terms of material aspects (Table 5), construction aspects (Table 6), and language aspects (Table 7).

Table 5. Validation Results of BM-3T Material Aspects

Test Items	Validator			Average	Category
	I	II	III		
1	3.75	3.75	3.25	3.58	Very Valid
2	3.75	3.75	3.25	3.58	Very Valid
3	3.75	3.75	3.25	3.58	Very Valid
4	3.75	3.50	3.00	3.42	Valid
5	3.75	3.50	3.00	3.42	Valid
6	3.75	3.50	3.00	3.42	Valid
7	3.75	4.00	3.00	3.58	Very Valid
8	3.75	3.25	3.25	3.42	Valid
9	3.75	3.00	3.25	3.33	Valid
10	3.75	3.25	3.50	3.50	Valid
11	4.00	3.25	3.00	3.42	Valid
12	4.00	3.50	3.50	3.67	Very Valid
13	4.00	3.00	3.00	3.33	Valid
14	4.00	3.50	3.50	3.67	Very Valid
15	4.00	3.75	3.00	3.58	Very Valid
16	3.75	3.25	3.00	3.33	Valid
Average Score	3.83	3.47	3.17	3.49	Valid

Assessment of the validity of the material obtained a score of 3.49. It is included in the valid category with details of 9 items in the valid category and seven items in the very valid category. Thus, we declared the test items developed to have met the material validity requirements to be feasible to use. The decision on the validity category of this test is following the score interpretation criteria in Table 2 that the test instrument is declared valid if it reaches a minimum score of 2.51. It is shown that the content of the developed test follows the subject's learning outcomes. Arikunto (2015) states that the validity of the material or content is the validity seen in terms of the content of the test, whether the contents represent representatively the entire material or subject matter being tested.

The construction validity assessment obtained a score of 3.42. It is means in the valid category with details of all test items in the valid category. Thus, the developed test items have met the construction validity requirements to be feasible. The decision on the validity category of this test is following the score interpretation criteria in Table 2 that the test instrument is declared valid if it reaches a minimum score of 2.51. So, we developed the tests to follow the correct aspects of thinking and question construction.

Table 6. Validation Results of BM-3T Construction Aspects

Test Items	Validator		Average	Category	
	I	II	III		
1	3.75	3.50	3.25	3.50	Valid
2	3.75	3.50	3.25	3.50	Valid
3	3.75	3.50	3.25	3.50	Valid
4	3.78	3.25	3.22	3.42	Valid
5	3.75	3.25	3.25	3.42	Valid
6	3.75	3.25	3.25	3.42	Valid
7	3.88	3.13	3.00	3.34	Valid
8	3.88	3.50	3.00	3.46	Valid
9	3.88	3.50	3.00	3.46	Valid
10	3.63	3.63	3.00	3.42	Valid
11	3.50	3.38	3.00	3.29	Valid
12	3.67	3.38	3.00	3.35	Valid
13	3.88	3.25	3.00	3.38	Valid
14	3.88	3.63	3.00	3.50	Valid
15	3.63	3.50	3.00	3.38	Valid
16	3.88	3.38	3.00	3.42	Valid
Average Score	3.77	3.41	3.09	3.42	Valid

Table 7. Results of BM-3T Language Aspect Validation

Test Items	Validator			Average	Category
	I	II	III		
1	4.00	4.00	3.50	3.83	Very Valid
2	4.00	4.00	3.50	3.83	Very Valid
3	4.00	4.00	3.50	3.83	Very Valid
4	4.00	4.00	3.50	3.83	Very Valid
5	4.00	4.00	3.50	3.83	Very Valid
6	4.00	4.00	3.50	3.83	Very Valid
7	4.00	3.25	3.75	3.67	Very Valid
8	4.00	3.50	3.75	3.75	Very Valid
9	4.00	3.75	3.75	3.83	Very Valid
10	4.00	3.25	3.75	3.67	Very Valid
11	4.00	3.75	3.75	3.83	Very Valid
12	4.00	3.50	3.00	3.50	Valid
13	4.00	3.75	3.00	3.58	Very Valid
14	4.00	3.25	3.00	3.42	Valid
15	4.00	3.75	4.00	3.92	Very Valid
16	4.00	3.75	3.75	3.83	Very Valid
Average Score	4.00	3.72	3.53	3.75	Very Valid

Assessment of the validity of the language obtained a score of 3.75, so it is included in the very valid category. There are 2 test items in the valid category and 14 test items in the very valid category. Thus, we declared the test items developed to have met the language validity requirements to be feasible to use. The decision on the validity category of this test is following the score interpretation criteria in Table 2 that the test instrument is declared valid if it reaches a minimum score of 2.51. The developed test has met the requirements of straightforwardness in delivering the message.

Based on the validation results, we can see that we rejected none of the test items. However, the validator scores 2 (disagree) in response to certain parts of the test items that are considered insufficient for improvement. The test items received a poor rating because there were still images that were not clear, and there were spelling errors. The final BM-3T has been revised according to feedback by validators.

Reliability is the consistency of measurement, namely how consistent test scores or assessment results are from one measure to another (Arikunto, 2014). The calculation of the reliexamability coefficient in this study is based on the test trial results using the KR-20 formula. The recap of the *four-tier multiple-choice test* reliability calculation can be seen in Table 8.

Table 8. Recap of Reliability Calculation of BM-3T

Test Items	p_i	q_i	Calculation	
1	0.43	0.57	Total Variance	Reliability
2	0.64	0.36	$S_t^2 = \frac{\sum X_t^2 - \frac{(\sum X_t)^2}{N}}{N}$	$r_i = \frac{k}{k-1}\left\{\frac{S_t^2 - \sum p_i q_i}{S_t^2}\right\}$
3	0.79	0.21	$S_t^2 = \frac{887 - \frac{(99)^2}{14}}{14}$	$= \frac{16}{16-1}\left\{\frac{13,35 - 3,57}{13,35}\right\}$
4	0.43	0.57		$= 1,07\left\{\frac{9,78}{13,35}\right\}$
5	0.50	0.50	$= \frac{887 - 700,07}{14}$	= 1,07 × 0,733
6	0.57	0.43	$= \frac{186,93}{14}$	= 0,78
7	0.29	0.71	= 13,35	
8	0.50	0.50		
9	0.50	0.50		
10	0.36	0.64		
11	0.29	0.71		
12	0.57	0.43		
13	0.21	0.79		
14	0.14	0.86		
15	0.57	0.43		
16	0.29	0.71		
k = 16; N = 14; $\sum X_t = 99$; $\sum X_t^2 = 887$; $\sum p_i q_i = 3,57$				$0,80 \le r_i \le 1,00$ $0,80 \le (0,78) \le 1,00 \rightarrow$ Reliabel

The calculation results show the reliability coefficient value of 0.78, so the test is declared reliable. Following the criteria for interpreting the degree of test reliability in Table 3, the test instrument is said reliable if it reaches a minimum reliability coefficient of 0.60 (Arikunto, 2014). Thus, BM-3T has stable characteristics to identify the misconceptions of prospective science teachers on the topic of biological macromolecules.

We analyzed Sixteen test items on the BM-3T t determine the difficulty level. Table 9 shows the summary of the difficulty level analysis calculation.

Table 9. Results of Analysis of the Difficulty Level of Test Items on the BM-3T

Test Items	Percentage	Category
-	-	Easy
1, 2, 3, 4, 5, 6, 8, 9, 10, 12, 15	68.75%	Moderate
7, 11, 13, 14, 16	31.25%	Difficult

The results of data analysis showed that a total of 16 test items were developed spread out only into two categories of test item difficulty level. None of the test items falls into the easy category. There are 11 test items in the moderate category with 68.75%, including test items numbered 1, 2, 3, 4, 5, 6, 8, 9, 10, 12, and 15. On the other hand, there are 5 test items in the difficult category with a percentage of 31.25%, namely test items numbered 7, 11, 13, 14, 16. In addition, calculation of the level of difficulty of BM-3T shows an average value of 0.44 so that it is included in the moderate category.

Conclusion

This research has succeeded in developing a diagnostic test in the form of a three-tier test named the Biological Macromolecules Three-Tier Test (BM-3T). The research results and data analysis show that the BM-3T is valid and reliable. Also, the level of difficulty of BM-3T shows an average value of 0.44, which is included in the moderate category. Therefore, BM-3T is appropriate to identify misconceptions of prospective science teachers on the topic of biological macromolecules.

Acknowledgements

We thank the Research and Community Service Institute at the Universitas Negeri Manado for funding this research in the 2022 Leading Higher Education Applied Research scheme with the number 694/UN41.9/TU/2022. Thanks to all validators who have validated and provided input on the development of the BM-3T. Also, thanks to all prospective science teachers (undergraduate science education students) willing to participate in this research.

References

Adityawardani, D. & Hidayati, S. N. (2017). Profil konsepsi siswa SMP dengan *CRI test* berbasis *revised Bloom's taxonomy* pada materi klasifikasi materi dan perubahannya. *E-Jurnal Pensa, 5*(3), 335-340. Retrieved from: https://ejournal.unesa.ac.id/index.php/pensa/article/view/21423

Al-Rsa'I, M. S., Khoshman, J. M., & Abu Tayeh, K. (2020). Jordanian Pre-Service Physics Teacher'smisconceptions about Force and Motion. *Journal of Turkish Science Education, 17*(4), 528-543. https://doi.org/10.36681/tused.2020.43

Arikunto, S. (2014). *Prosedur Penelitian: Suatu Pendekatan Praktik*. Jakarta: Rineka Cipta.

Arikunto, S. (2015). *Dasar-dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara.

Arslan, H. O., Cigdemoglu, C., & Moseley, C. (2012). A three-tier diagnostic test to assess pre-service teachers' misconceptions about global warming,

greenhouse effect, ozone layer depletion, and acid rain. *International Journal of Science Education, 34*(11), 1667–1686. https://doi.org/10.1080/09500693.2012.680618

Boodhun, N. (2018). Seeing is believing: structures and functions of biological molecules. *BioTechniques, 64*(4), 143–146. https://doi.org/10.2144/btn-2017-0123

Çam, A., Topçu, M. S., & Sülün, Y. (2015). Preservice science teachers' attitudes towards chemistry and misconceptions about chemical kinetics. *Asia-Pacific Forum on Science Learning and Teaching, 16(2), 1-16.* Retrieved from: https://www.eduhk.hk/apfslt/download/v16_issue2_files/cam.pdf

Fatokun, K. V. F. (2016). Instructional misconceptions of prospective chemistry teachers in chemical bonding. *International Journal of Science and Technology Educational Research, 7*(2), 18-24. https://doi.org/10.5897/IJSTER2016.0357

Gronlund, N. E. (1993). *How to Make Achievement Tests and Assessments*. Boston, USA: Allyn and Bacon.

Gurel, D. K., 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. https://doi.org/10.12973/eurasia.2015.1369a

Jaswal, S. S., O'Hara, P. B., Williamson, P. L., & Springer, A. L. (2013). Teaching structure: Student use of software tools for understanding macromolecular structure in an undergraduate biochemistry course. *Biochemistry and Molecular Biology Education, 41*(5), 351-359. https://doi.org/10.1002/bmb.20718

Juliani, Yuzrial, & Huda, I. (2021). Development of four-tier multiple choice diagnostic test to know students' misconceptions in science learning. *Jurnal Penelitian Pendidikan IPA, 7*(4), 763-769. https://doi.org/10.29303/jppipa.v7i4.854

Jusniar, J., Effendy, E., Budiasih, E., & Sutrisno, S. (2020). Developing a three-tier diagnostic instrument on chemical equilibrium (TT-DICE). *Educación Química, 31*(3), 84-102. http://dx.doi.org/10.22201/fq.18708404e.2020.3.72133

Kanli, U. (2014). A study on identifying the misconceptions of pre-service and in-service teachers about basic astronomy concepts. *Eurasia Journal of Mathematics, Science and Technology Education, 10*(5), 471-479. https://doi.org/10.12973/eurasia.2014.1120a

Karakaya, F., Yilmaz, M., & Aka, E. I. (2021). Examination of pre-service science teachers' conceptual perceptions and misconceptions about photosynthesis. *Pedagogical Research, 6*(4), 1-11. https://doi.org/10.29333/pr/11216

Karataş, A. (2020). Preservice science teachers' misconceptions about evolution. *Journal of Education and Training Studies, 8*(2), 38-46. https://doi.org/10.11114/jets.v8i2.4690

Karim, S. A., Sudiro, S., & Sakinah, S. (2021). Utilizing test items analysis to examine the level of difficulty and discriminating power in a teacher-made test. *EduLite: Journal of English Education, Literature, and Culture, 6*(2), 256-269. http://dx.doi.org/10.30659/e.6.2.256-269

Katalog Jurusan Pendidikan IPA FMIPA Universitas Negeri Manado. (2021).

Kurtulus, M. A. & Tatar, N. (2021). An analysis of scientific articles on science misconception: A bibliometric research. *Elementary Education Online, 20*(1), 192-207. https://doi.org/10.17051/ilkonline.2021.01.022

Liampa, V., Malandrakis, G. N., Papadopoulou, P., & Pnevmatikos, D. (2019). Development and evaluation of a three-tier diagnostic test to assess undergraduate primary tachers' understanding of ecological footprint. *Research in Science Education, 49*(3), 711-736. https://doi.org/10.1007/s11165-017-9643-1

Mason, K. A., Losos, J. B., & Singer, S. R. (2016). *Biology 11th Edition*. New York: McGraw-Hill Education.

Ozkan, G. & Ackay, H. (2016). Preservice science teachers' beliefs about astronomy concepts. *Universal Journal of Education Research, 4*(9), 2092-2099. https://doi.org/10.13189/ujer.2016.040921

Pesman, H. & Eryilmaz, A. (2010). Development of a three-tier test to assess misconceptions about simple electric circuits. *The Journal of Educational Research, 103*(3), 208-222. https://doi.org/10.1080/00220670903383002

Riduwan (2012). *Skala Pengukuran Variabel-Variabel Penelitian*. Bandung: Alfabeta.

Safo-Adu, G. (2020). Remediating pre-service integrated science teachers' misconceptions about acid-base concepts using cognitive conflict instructional strategy. *American Journal of Education and Information Technology, 4*(2), 86-98. https://doi.org/10.11648/j.ajeit.20200402.17

Soeharto, S., 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

Susanti, R. (2018). Misconception of biology education student of teacher training and education of Sriwijaya University to the concept of photosynthesis and respiration. *IOP Conf. Series: Journal of Physics: Conf. Series, 1022*(012056), 1-7. https://doi.org/10.1088/1742-6596/1022/1/012056

Suwarto. (2013). *Pengembangan Tes Diagnostik Dalam Pembelajaran*. Yogyakarta: Pustaka Pelajar.

Türkogus, S. (2020). Investigation of three-tier diagnostic and multiple choice test on chemistry concepts with response change behavior. *International Education Studies, 13*(9), 10-22. https://doi.org/10.5539/ies.v13n9p10

Wahyuni, A. S. A., Rustaman, N., Rusdiana, D., & Muslim. (2019). Paper conceptions and misconceptions of pre-service teacher about light. *Proceedings of the 1st International Conference on Advanced Multidisciplinary Research (ICAMR 2018),* Makassar – Indonesia, 56-61.

Widiarti, H. R., Marfuaf, S., & Retnosari, R. (2019). Identifying students' misconception about intermolecular forces topic in organic chemistry I course. *Unnes Science Education Journal, 8*(1), 46-56. https://doi.org/10.15294/usej.v8i1.23527

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. https://doi.org/10.21070/sej.v4i2.878

Yang, DC. & Sianturi, I. A. J. (2019). Assessing students' conceptual understanding using an online three-tier diagnostic test. *Journal of Computer Assisted Learning, 35*(5), 1-12. https://doi.org/10.1111/jcal.12368

Yangin, S., Sidekli, S., & Gokbulut, Y. (2014). Prospective teachers' misconceptions about classification of plants and changes in their misconceptions during pre-service education. *Journal of Baltic Science Education, 13*(1), 106-117. Retrieved from: http://oaji.net/articles/2015/987-1437062464.pdf

Zaman, A., Niwaz, A., Faize, F. A., Dahar, M. A., & Alamgir. (2010). Analysis of multiple choice items' sequencing on difficulty level in the test of mathematics. *European Journal of Social Sciences, 17*(1), 61-67. Retrieved from: https://www.researchgate.net/publication/261367711_Analysis_of_... Mathematics

Zuhdi, M. & Busyairi, A. (2021). Misconception Profile of Prospective Physics Teachers Assessed from Various Representations of Electricity Subjects. *Journal of Science and Science Education, 2*(1), 33-37. https://doi.org/10.29303/jossed.v2i1.490