

PAPER • OPEN ACCESS

Analysis of higher order thinking skills instrument test for pre-service biology teachers based on marine ecology toward sustainable development

To cite this article: J Jumrodah *et al* 2021 *J. Phys.: Conf. Ser.* **1731** 012011

View the [article online](#) for updates and enhancements.

You may also like

- [The Content Validity: Two-Tier Multiple Choices Instrument to Measure Higher-Order Thinking Skills](#)
S R Maulita, Sukarmin and A Marzuki
- [Development of two-tier multiple choice instrument to measure science process skill](#)
N A A Sholihah, Sarwanto and N S Aminah
- [The development of an instruments diagnostic two-tier multiple choice assisted smartphone on salt hydrolysis](#)
R A Bachtiar, N Nahadi and H Firman



245th ECS Meeting
San Francisco, CA
May 26–30, 2024

PRiME 2024
Honolulu, Hawaii
October 6–11, 2024

Bringing together industry, researchers, and government across 50 symposia in electrochemistry and solid state science and technology

Learn more about ECS Meetings at
<http://www.electrochem.org/upcoming-meetings>

ECS Save the Dates for future ECS Meetings!

Analysis of higher order thinking skills instrument test for pre-service biology teachers based on marine ecology toward sustainable development

J Jumrodah^{1,2,*}, S Liliarsari¹, Y H Adisendjaja¹ and Y Sanjaya¹

¹ Program Studi Pendidikan IPA, Sekolah Pascasarjana, Universitas Pendidikan Indonesia Jl. Dr. Setiabudhi No. 229, Bandung 40154, Indonesia

² Program Studi Tadris Biologi, Institut Agama Islam Negeri Palangka Raya, Jl. G. Obos Komplek Islamic Centre, Palangka Raya, Indonesia

*jumrodah@iain-palangkaraya.ac.id

Abstract. This research used quantitative descriptive approach. It aims to analysis the quality of instruments test include validity, reliability, difficulty level and index of determinant to develop tests of critical thinking skills and creative thinking skills of prospective pre-service biology teachers on marine ecology theory. The test used instrument consisted of 2 type namely question descriptions and two tier multiple choice. There were two test of validation, namely content validation and empirical validation. The results of content validation concluded the question tests that feasible to be used with improvements. The results of empirical validation show that the ten items are valid question descriptions category and two invalid categorized questions. Where as instrument test twenty-three of two tier multiple choice valid in categories and four invalid categorical questions. Question descriptions instrument test used reliability analysis using Cronbach Alpha $r_{11} = 0.73$ (high). Instrument test two tier multiple choice used KR-20 $r_{11} = 0.83$ (very high), therefore the instruments test made are suitable for implementation in marine ecology. The average value of high order thinking skills for the question descriptions instrument test achieved by students was 26, while the two tier multiple choices instrument test 28. The highest value of the description questions instrument test achieved by students is 56 and the lowest score is 3. The highest value of two tier multiple choice questions instrument test achieved by students is 56 and the lowest score is 11. It shows that the ability of high order thinking skills of students is still lacking.

1. Introduction

Based on the results of curriculum studies at Universitas Hasanudin Makassar, marine ecology material not only discusses the concept of the material but also must be able to criticize in the analysis of marine biological resources, for example by cultivation marine biota specifically seaweed, kima clams and sea horse [1]. Some universities have integrated scientific knowledge and attitudes as a basis for problems solving in the marine environment, but specific skills such as producing projects by developing marine organisms. It also producing marine potential that are sustainable and evaluating them have not yet been carried out. It is deemed necessary to equip critical thinking skills and creative thinking skills in learning marine ecology towards sustainable development. In the implementation it is necessary used test instruments of critical thinking skills and creative thinking skills.



One of the goals of 21st century education is to prepare students to understand in-depth knowledge and apply effective critical thinking skills to overcome challenges in an ever-changing society [2, 3]. Critical thinking needs to be mastered before reaching creative thinking, problem solving and decision making [4]. In addition, learning should be conditioned on real situations to analyze problems and making decisions in developing critical thinking skills and creative thinking skills, because these skills are very important and should be possessed by pre-service biology teachers [5].

Creative thinking is the ability to feel the lack of an object, shape and test new hypotheses, then communicate the results obtained [6]. Divergent thinking skills are shown by fluency, flexibility, unique thinking, sensitivity to problems and reshaping existing ideas [7]. Creative thinking skills can develop the skills of discovery towards original, aesthetic and constructive ideas. It is related to perspectives, concepts, emphasize intuitive and rational thinking, especially in the use of information and materials to express or explain with the original perspectives of thinkers [8]. Indicators of creative thinking used in this instrument referred to Torrance framework including: flexibility, originality and elaboration.

Previous research that has been conducted to the development of higher-order thinking skills in marine ecology learning such as the implementation of problem-based learning [9] and recent innovations in marine biology [10]. The aim of this research is to analyze the higher order thinking skills test for pre-service biology teachers based on marine ecology towards sustainable development.

2. Materials and methods

This research uses a quantitative descriptive approach. Involving 36 students from semesters 7 and 9. Research subjects were taken with the consideration that those pre-service biology teacher candidates had received marine ecology learning. In addition, the number of 7 semester students is not up to 20, so it is added in 9 semesters. The test instrument consisted of 2 types of tests, namely 27 multiple choice items and 12 questions description of test items. The higher-order thinking skills developed in this instrument are critical thinking skills and creative thinking skills. In practice, this study aims to analyze the quality of test questions including validity, reliability, difficulty index and discriminatory strength. Instrument analysis using SPSS and ANA-test. The test instrument consisted of 2 types of tests, they were 27 items two tier multiple choice and 12 items question descriptions tests. The higher order thinking skills developed in this instrument were critical thinking skills and creative thinking skills. In practice, this study aims to analyze the quality of the test questions including validity, reliability, difficulty index and discriminating power. The analysis of the instruments were used SPSS and ANA-test.

3. Results and discussion

3.1. Analysis of test instruments

Before validating the instrument empirically, the content was validated by three experts (2 lecturers from Universitas Pendidikan Indonesia and 1 lecturer from Universitas Palangkaraya). The goal is to get suggestions and improvements to the test items that will be used. The results of the content validation from expert judgment can be concluded that test questions are appropriate to use with several improvements. It is including the suitability of items with the learning objectives with indicators of critical thinking skills and creative thinking skills. Then, the test items were revised based on expert advice. Furthermore, empirical validation was conducted with the subject of pre-service biology teachers from one of the universities in Bandung. Students are given 60 minutes to explain of 12 question descriptions and 81 minutes to work test two tier multiple choice. The results of the analysis question descriptions are shown in Table 1 and the results of the analysis of two tier multiple choice test in Table 2.

Table 1. Results of empirical validation analysis on question descriptions.

Item test	Validity		Reliability		Difficulty Index		Discriminating Power		Explanation
	Score	Intpretation	Score	Intpretation	Score	Intpretation	Score	Intpretation	
1	0,271	Invalid			0,52	Medium	0,10	Bad	Not used
2	0,499	Validity			0,57	Medium	0,40	Very Good	Used
3	0,527	Validity			0,24	Difficult	0,20	Sufficient	Used
4	0,572	Validity			0,17	Difficult	0,40	Very Good	Used
5	0,533	Validity			0,21	Difficult	0,27	Sufficient	Used
6	0,621	Validity	0,73	Reliable	0,21	Difficult	0,37	Sufficient	Used
7	0,569	Validity			0,41	Medium	0,40	Very Good	Used
8	0,561	Validity			0,17	Difficult	0,30	Sufficient	Used
9	0,236	Invalid			0,10	Difficult	0,10	Bad	Not Used
10	0,545	Validity			0,11	Difficult	0,23	Sufficient	Used
11	0,639	Validity			0,18	Difficult	0,37	Sufficient	Used
12	0,533	Validity			0,300	Medium	0,33	Sufficient	Used

Based on Table 1 it is known that there are 2 invalid questions, namely number 1 and 9. For the discriminating power 2 question bad category, while the difficulty index is the number 1 question while the number 2 question is difficult. So that the two questions cannot be used as further instruments.

Table 2. Results of empirical validation analysis on two tier multiple choice.

Item test	Validity		Reliability		Difficulty Index		Discriminating Power		Explanation
	Score	Intpretation	Score	Intpretation	Score	Intpretation	Score	Intpretation	
1	0,583	Validity			0,130	Difficult	0,40	Very Good	Used
2	0,597	Validity			0,176	Difficult	0,53	Very Good	Used
3	0,487	Validity			0,426	Medium	0,40	Very Good	Used
4	0,396	Validity			0,231	Difficult	0,27	Medium	Used
5	0,461	Validity			0,528	Medium	0,50	Very Good	Used
6	0,479	Validity			0,370	Medium	0,47	Very Good	Used
7	0,464	Validity			0,426	Medium	0,40	Very Good	Used
8	-0,036	Invalid			0,361	Medium	0,03	Bad	Not Used
9	0,159	Invalid			0,093	Difficult	0,13	Bad	Not Used
10	0,628	Validity			0,287	Difficult	0,70	Very Good	Used
11	0,616	Validity			0,306	Medium	0,37	Good	Used
12	0,424	Validity			0,546	Medium	0,73	Very Good	Used
13	0,183	Invalid	0,83	Reliable	0,130	Difficult	0,10	Bad	Not Used
14	0,000	Invalid			0,000	Difficult	0,00	Bad	Not Used
15	0,578	Validity			0,130	Difficult	0,27	Medium	Used
16	0,533	Validity			0,222	Difficult	0,30	Good	Used
17	0,374	Validity			0,194	Difficult	0,23	Medium	Used
18	0,555	Validity			0,194	Difficult	0,40	Very Good	Used
19	0,431	Validity			0,157	Difficult	0,20	Medium	Used
20	0,341	Validity			0,185	Difficult	0,23	Medium	Used
21	0,479	Validity			0,370	Medium	0,43	Very Good	Used
22	0,599	Validity			0,111	Difficult	0,27	Medium	Used
23	0,516	Validity			0,139	Difficult	0,30	Good	Used
24	0,347	Validity			0,296	Difficult	0,23	Medium	Used
25	0,614	Validity			0,509	Medium	0,43	Very Good	Used
26	0,515	Validity			0,481	Medium	0,40	Very Good	Used
27	0,420	Validity			0,509	Medium	0,40	Very Good	Used

In Table 2. there are 4 invalid questions, they are questions 8, 9, 13 and 14. The discriminating power of the four questions has a bad category. For the difficulty index found in test number 8 with the medium category, while for all three questions the difficult category. So that the four questions cannot be used as further instruments.

Based on the analysis result of the two types of questions tested, there are 6 invalid questions. This means that the six items cannot measure the critical thinking and creative thinking abilities possessed by students. Validity is a degree of measurement that reflects the expected content domain, so validity is important for assessment test [11]. Testing the validity of test questions is also determined by the number of samples the more samples are used the more valid test questions. In this study using sample 36, so it has exceeded the minimum limit. Additional validation studies that involve a larger and diverse

group of respondents representing the target population should be conducted to further strengthen the quantitative data set and related measures [12].

The validity and reliability of good items are influenced by several factors. There are four factors that cause good validity and reliability [13], namely (a) items are developed in accordance with development procedures, (b) items are developed from appropriate references, (c) items are passed through the stages content validation, and (d) items were empirically tested with respondents who worked hard and were closely watched. All of these factors have been carried out in this study, so the test questions in this study have good validity and reliability.

Based on an analysis of student answers, it is known there are 6 discriminating power question with the poor category. This is because the quality of the questions is not able to distinguish students from low groups with high groups. For example, in question number 13 students from the low group are able to answer while students from the high group cannot to answer. The discriminating power is a measurement to compare the number of people with high test scores who answered that item correctly with the number of people with low scores who answered the same item correctly [13]. The higher the discriminating power, the better the item because such a value indicates that the item discriminates in favor of the upper group, which should get more items correct [14].

Based on the analysis result of the two types of tests tested, it was found that 62% of difficult questions almost all students could not answer, both from the achiever group and the low group. This is because marine biota cultivation material has never been obtained before, but before the student test is conducted, it has been given articles related to marine biota cultivation. The purpose of analyzing the level of difficulty of the questions is to determine the quality of good questions. Test instruments must have easy, medium and difficult question classifications [15].

After finding out that there were 10 question descriptions items that were declared valid. 23 two tier multiple choice items were valid. Reliability test was conducted to determine the level of diversity, when used to measure the ability of critical thinking skills and creative thinking skills of prospective pre-service biology teachers. Based on the calculation results, the reliability coefficient values obtained on the description used Cronbach Alpha $r_{11} = 0.73$, which means the item has a high level of reliability, while the value of the reliability coefficient on the two tier multiple choice used the KR-20 $r_{11} = 0.83$ which means the item has very high level of reliability. Reliability of a question is a requirement on a test as an evaluation tool so that good questions are questions that have high reliability [16].

3.2. Analysis of student test results

Pre-Service Biology Teachers high-level thinking ability on marine ecology material towards sustainable finding can be seen in Table 3.

Table 3. Test results for higher level thinking.

Number	Test type	Aspect	Explanation
1	description test	Mean score	26
		Standard Deviation	12
		The highest score	56
		Low score	3
		Maximum Score	100
2	MC two-tier test	Mean score	28
		Standard Deviation	13
		The highest score	56
		Low score	11
		Maximum Score	100

Based on the results in Table 3, it is known that the higher order thinking ability of Biology Education students in one of the universities in Bandung is still lack. This is evidenced by the average value of students' high thinking skills for the question descriptions items only at 26 and the two tier multiple

choice question is 28. These results are the same as the results of the study of which show students' critical thinking skills at a value of 27.2 [17]. The research results of Pradana et al. [18] there are 74% students have very low categories in critical thinking skills, because of the inability to answer questions related to the consideration of whether the source is reliable, make deductions, identifies assumptions and determines of an action, so it can be concluded that the critical thinking skills of Biology pre-service teachers are in very low category [5]. The profile of pre-service physics teachers (PPTs)' thinking styles and critical thinking skills can be categorized as low [19]. This further reinforces that students' critical thinking skills, including students in Indonesia, are still lack. This has made Indonesian students less able to compete internationally [20].

4. Conclusion

Based on the results of research conducted test instruments of higher order thinking skills in the marine ecology for pre-service biology teachers developed declared valid, reliable and feasible to be used. The instrument of higher order thinking skills test in this study is an instrument test two tier multiple choice and question descriptions tests. Higher order thinking skills developed are critical thinking skills with four indicators include: basic support, inference, advanced clarification, and strategies and tactics and the creative thinking skills included 3 indicators, include: flexibility, originality and elaboration. Therefore, tests that have been declared valid and reliable are then used to measure the ability of higher order thinking skills. The average value of high order thinking skills for the question descriptions instrument test achieved by students was 26, while the two tier multiple choices instrument test 28. The highest value of the description questions instrument test achieved by students is 56 and the lowest score is 3. The highest value of two tier multiple choice questions instrument test achieved by students is 56 and the lowest score is 11. It shows that the ability of high order thinking skills of students is still lacking. The researcher can re-develop the higher order thinking skills test instrument by adjusting the performance indicators to the question. It will develop marine ecology learning with the project approach.

Acknowledgments

The author would like to thank the Pre-Service Biology Teachers Biology Education Study Program at one of the universities in Bandung who will to be the sample of this study. The author also wishes to thank the lecturers who provided the opportunity to conduct this research.

References

- [1] Syafriman A, Hayati and Feby A R 2018 Analisis Kesesuaian Parameter Lingkungan Untuk Kegiatan Rancing Kuda Laut (*Hyppocampus sp*) di Perairan Pulau Barrang Caddi *Hasanuddin Student Journal* **2** 1 231-238
- [2] Atabaki A M S, Keshtiaray N and Yarmohammadian M H 2015 Scrutiny of Critical Thinking Concept *International Education Studies* **8** 3
- [3] Arslan R, Gulveren H and Aydin E 2014 A research on critical thinking tendencies and factors that affect critical thinking of higher education students *International Journal of Business and Management* **9** 5 43
- [4] L Liliyasi 2008 *Berpikir Kritis dalam Pembelajaran Sains Kimia Menuju Profesionalitas Guru*
- [5] Jumrodah J, Liliyasi S and Adisendjaja Y H 2019 Profile of pre-service biology teachers critical thinking skills based on learning project toward sustainable development *Journal of Physics: Conference Series* **1157** 2
- [6] Bockers A, Mayer C and Bockers T M 2014 Does learning in clinical context in anatomical sciences improve examination results, learning motivation, or learning orientation?. *Anatomical sciences education* **7** 1 3-11
- [7] Wulandari N A D and Sukestiyarno S 2017 Development of OQALE Based Reference Module for School Geometry Subject and Analysis of Mathematical Creative Thinking Skills *Journal of Physics: Conference Series. Conf. Series* **824** 012048

- [8] Tawil M and Liliyasi L 2013 *Berpikir Kompleks dan Implementasinya dalam Pembelajaran IPA* (Makassar: Badan Penerbit Universitas Negeri Makasar)
- [9] Magsino R M 2014 Enhancing higher order thinking skills in a marine biology class through Problem-Based Learning *Asia Pacific Journal of Multidisciplinary Research* **2** 5
- [10] Boero F 2009 Recent innovations in marine biology *Marine Ecology* **30** 1-12
- [11] Istiyono E, Mardapi D and Suparno 2014 Pengembangan tes kemampuan berpikir tingkat tinggi Fisika (PysTHOTS) peserta didik SMA *Jurnal Penelitian dan Evaluasi Pendidikan* **18** 11-12
- [12] Tiruneh D T, Cock M D, Weldeslassie A G, Elen J and Janssen R 2016 Measuring Critical Thinking in Physics: Development and Validation of a Critical Thinking Test in Electricity and Magnetism *International Journal of Science and Mathematics Education* 1-20
- [13] Sya'bandari Y, Firman H and Rusyat L 2017 The Development and Validation of Science Virtual Test to Assess 7th Grade Students' Critical Thinking on Matter and Heat Topic *Journal of Science Learning* **1** 1 17-27
- [14] Hetzel S M 1997 Basic Concepts in Item and Test Analysis *Clearinghouse on Assessment and Evaluation* 1-24
- [15] Kadir A 2015 Menyusun dan Menganalisis Tes Hasil Belajar *Al-Ta'dib* **8** 2 70-81
- [16] Amalia A N and Widayati A 2012 Analisis butir soal tes kendali mutu kelas XII SMA mata pelajaran ekonomi akuntansi di kota Yogyakarta tahun 2012 *Jurnal Pendidikan Akuntansi Indonesia* **10** 1
- [17] Pradana S D S, Parno P and Handayanto S K 2017 Pengembangan tes kemampuan berpikir kritis pada materi Optik Geometri untuk mahasiswa Fisika *Jurnal Penelitian dan Evaluasi Pendidikan* **21** 1 51-64
- [18] Pradana S D S, Parno P and Handayanto S K 2016 Kemampuan Berpikir Kritis Mahasiswa Tahun Pertama Jurusan Fisika Universitas Negeri Malang *Seminar Nasional Pendidikan IPA, Pascasarjana Universitas Negeri Malang Malang* 8 Oktober 2016
- [19] Saprudin S, Liliyasi S, Prihatmanto A S and Setiawan A 2019 Pre-service physics teachers' thinking styles and its relationship with critical thinking skills on learning interference and diffraction *Journal of Physics: Conference Series* **1157** 3 032029
- [20] Frijters S, ten Dam G and Rijlaarsdam G 2008 Effects of dialogic learning on value-loaded critical thinking *Learning and Instruction* **18** 1 66-82