Jurnal Studi Guru dan Pembelajaran, Vol. 4, No. 1, April 2021

Assessment to Measure Problem Solving Ability in the Lesson of the Interaction of Living Things with the Environment

Raden Gamal Tamrin Kusumah[1], Rambat Nur Sasongko[2], Muhammad Kristiawan[3], Ahmad Walid[4], Meri Andaria[5]

[1,4] Institut Agama Islam Negeri Bengkulu, Indonesia
[2, 3, 5] Doktor Pendidikan Universitas Bengkulu, Indonesia

* raden@iainbengkulu.ac.id

Abstract

This study aims to determine the form of assessment development to measure students' problemsolving abilities. This research uses a type of Research and Development (R & D) with a quantitative approach. The research procedure refers to the Borg & Gall development model, which includes 8 steps, namely potential problems, information gathering, product development, expert validation, expert revision, small group trials, product revision, final product. The research instruments used were questionnaires, interviews, and documentation. The research was conducted in 3 schools in Bengkulu City. The resulting assessment has categories, namely defining problems, examining problems, planning solutions, implementing plans that have been made, evaluating. After carrying out various kinds of expert tests, this question instrument has been valid and reliable, so it is appropriate to be disseminated to the school community.

Keywords: *assessment. Measure problem solving ability, Interaction of Living Things with Environment*

Introduction

The development of science and technology in the 21st century provides new challenges in the world of education (Putra et al., 2018; Rusmansyah et al., 2019; Trisnawati & Sari, 2019). Indonesia has experienced ten changes to the educational curriculum (Abdullah, 2007; Wirianto, 2014). The latest education curriculum, namely the 2013 curriculum (K13), will take effect from 2014 to 2021 (now). The essence of the current curriculum prioritizes the strengthening of character education, literacy, 21st century skills, and higher order thinking skills (HOTS) (Kusuma et al., 2017; Retnawati et al., 2018; Tanujaya et al., 2017; Thompson, 2008; Widana, 2018).

There are three mechanisms needed to impart 21st century skills, as the 2013 curriculum is implemented.The mechanisms are 1) the wider community must realize the importance of 21st century skills as today's education 2), schools should have a new design for learning about how people learn information processing, effective use of technology, and 21st century skills in an academic context 3), policy makers should contribute to formulating rules of assessment that can measure 21st century academic achievement and skills (Ellianawati et al., 2020).

21st century skills can be applied to science learning in schools. This is because science learning is very close to everyday life. Science material in science learning is very contextual to student life. Learning science is related to efforts to systematically understand various natural phenomena. In essence, science learning has four dimensions, namely attitude

https://doi.org/10.30605/jsgp.4.1.2021.552

(Alhunaini et al., 2020), process (Karamustafaoğlu, 2011), product (Cavanagh et al., 2016), and application (Kusumah et al., 2020). Attitude is related to curiosity about objects, natural phenomena, living things, and cause-and-effect relationships that cause new problems to be solved through correct procedures. Science learning is open ended. The process is related to problem-solving procedures using scientific methods which include formulating hypotheses, designing and carrying out investigations, collecting and analyzing data, and drawing conclusions. Science products in the form of exam-oriented learning include concepts, principles, laws, and theories. Application deals with the application of scientific methods and products in everyday life.

The dimensions of science learning that are applied in the city of Bengkulu have not been fully applied in learning in schools. This qualitative data was obtained from interviews with educators in junior high schools in Bengkulu. In general, teachers only provide routine questions that are the same as evaluation questions in textbooks. In addition, the assessment only measures aspects of remembering and understanding or cognitive level 1 in Bloom's taxonomy (C1). In fact, textbooks that support learning activities at school have provided various materials that can invite students to think actively and present various systematic material concepts (Astika et al., 2019; Dharmawati et al., 2016). However, in teacher assessment activities do not train students' thinking skills. Meanwhile, learning science requires an assessment that can train several thinking skills of students, one of which is the ability to solve problems. One of the science materials that are very close to student life is the material on the interaction between living things. Students when in the environment, both school and home environments, are close to nature in which there is an interaction between abiotic and biotic components. So that this extraordinary field experience needs to be discussed so that students are able to solve problems in their daily life.

To solve problems in student problem solving assessments, it is necessary to develop an assessment that can measure students' problem solving abilities on the material of interaction between living things and the environment.

Method

The procedure in this study used a product development model developed by Borg & Gall (1983). This research uses steps, namely identification of potential problems, gathering information, product development, expert validation, expert revision, small group trials, product revisions, and finally the final product dissemination. The stages of this research can be seen in Figure 1.

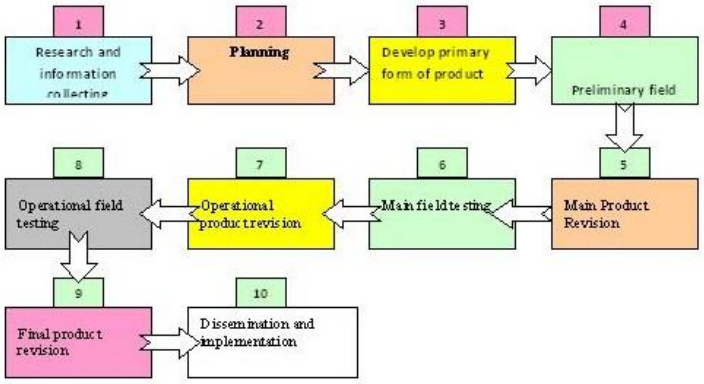

Figure 1. Borg & Gall's Research Procedures

Research involves experts being asked for their expertise to test products and validate products. The research and development process stages usually form a consistent cycle to produce a certain product according to needs, through the initial product design step, initial product testing to find various weaknesses, retrying, improving until finally a product that is considered ideal is found.

The types of data obtained in this assessment development research are qualitative data and quantitative data. Qualitative data were obtained from suggestions and corrective input from material experts and linguists. Quantitative data were obtained from student and teacher trials working on assessment instruments that were in the process of being developed. In addition, quantitative data were also obtained from initial trials of problem solving questions and final stage trials.

The expert test assessment sheet was made using an instrument whose results were made of the criteria listed in table 1

Table 1. Expert Validation Assessment Score

Information	Score
Very Good	5
Good	4
Enough	3
Less	2
very less	1

The validation results that have been listed in the validation sheet for the development of the assessment of the questions will be analyzed using the following formula:

$$N = \frac{k}{Nk} \times 100\%$$

The validity of the question instrument used the product moment correlation formula, namely :

$$r_{xy} = \frac{N\Sigma xy - (\Sigma x)(\Sigma y)}{\sqrt{\{N\Sigma x^2 - (\Sigma x)^2\}\{N\Sigma y^2 - (\Sigma y)^2\}}}$$

(Arikunto, 1997 : 146)

Based on the product moment correlation table with the provisions rcount> rtable. It means that the items are valid.

Furthermore, the percentage of the validation score interpretation criteria is based on table2

Table 2. Interpretation Criteria Score judgment

Interval Kriteria	Kriteria	Konversi
$86\% \leq N < 100\%$	Very Good	A
$72\% \leq N < 85\%$	Good	B
$58\% \leq N < 71\%$	Enough	C
$44\% \leq N < 57\%$	Less	D
$N \leq 44\%$	very less	E

Results and Discussion

Students' *problem solving* skills were analyzed from indicators that have been applied and modified from Mourtos et al., (2004). Indicators of *problem solving* skills can be seen in table 3.

Table 3. Modified Problem Solving Ability Indicator

No	Indicator	Indikator panel Component
1	Defining Problems	1.1 State facts related to the problem. 1.2 Define a concept or category. 1.3 Determine information / data related to the problem given
2	Inspect Problems	1.1 Identifying the root of the problem. 1.2 Checking the reciprocal relationship (cause-effect) of the given problem. 1.3 Checking the severity of the problem. 1.4 Checking the solutions that have been done to solve the related problem..
3	Planning The Solutions	1.5 Develop a problem-solving plan based on the root of the problem 1.6 Mapping sub-problems and sub-solutions. 1.7 Select theories, principles and approaches for solving related problems.
4	Conducted the Plans that've been made	1.8 List the problems to be resolved. 1.9 Sequencing the work steps related to the solutions that have been made. 1.10 Determine who needs to be contacted for information regarding the implementation of the solution.
5	Evalution	1.11 Checking the feasibility of the solutions created. 1.12 Making assumptions regarding the solutions being made. 1.13 Estimating the results that will be obtained through the solutions that have been made. 1.14 Choosing the right media, convey and communicate the solutions that have been made.

The initial stage for conducting assessment development begins, namely 1. Questionnaire for Analysis of Instrument Validation Results

Validation sheet containing statements. Then the validator fills out a questionnaire by giving a check mark on the categories provided by the researcher based on a Likert scale consisting of five assessment scores as follows:

Table 4. Expert Validation Assessment Score

Skor	Keterangan
5	Very Good (VG)
4	Good (G)
3	Enough (E)
2	Less (L)
1	very less (VL)

The validation results that are listed in the assessment validation sheet will be analyzed using the following formula:

$$\text{N}=\frac{K}{Nk} X\ 100\%$$

Legends:

N = Percentage of aspect eligibility

k = Score data collection results

Nk = maximum score (highest criterion score x number of aspects x number of validators)

Furthermore, the feasibility presentation obtained is then interpreted into categories based on the following table:

Table 5. Validation Score Interpretation Criteria

Criterium Intervali	Criterium	Konversi
86%≤N<100%	Sangat Baik	A
72%≤N<100%	Baik	B
58%≤N<100%	Cukup	C
44%≤N<100%	Kurang	D
N≤44%	Sangat Kurang	E

(Source: Sudjana, 2009)

Table 6. The results of the validation test can be displayed in the following table

Validator	Jumlah soal			Sum of instrumen test
	Valid	Kurang valid	Tidak valid	
Validator 1	13	1	1	15
Validator 2	11	2	2	15
Validator 3	13	1	2	15

Test the Validity of Question Items

Testing the validity of the items was carried out statistically using the Quest program. the validity of the items using the Quest program whose analysis is the Rasch model which can be seen with the goodness of fit, the overall fit test was developed by Adam and Khoo. The results of the validity test can be seen in table 7

Table 7. Results of instrument validity test

No Butir soal	r	Ket
1	0.25	Invalid
2	0.59	Valid
3	0.54	Valid
4	0.36	Valid
5	0.43	Valid
6	0.23	Invalid
7	0.15	Invalid
8	0.38	Valid
9	0.36	Valid
10	0.56	Valid
11	0.62	Valid
12	0.18	Invalid
13	0.70	Valid
14	0.38	Valid
15	0.32	Valid

R tabel 5% ; N 30 = 0.32

Based on the results of the validity test using the product moment correlation, it was obtained that the valid questions were questions number 2, 3, 4, 8, 9, 10, 11, 13, 14, 15 and the invalid questions were questions number 1, 6, 7. , and 12. The trigger for invalid questions in question number 1 was caused by the easy question so that many students responded with the correct answer, on the other hand the rest were questions number 6, 7 and 12 because few students answered the question correctly.

High validity gives legitimacy that this question instrument is suitable for use in both formative and summative exams for school students (Furnham, 2009; Huijgen et al., 2017; Life & Instrument, 2020; Rusmansyah et al., 2019).

Reliability

The reliability of the quest program can be seen in the summary of item estimates-reability of estimate. Reliability criteria are interpreted using the following references:

Table 8. Interpretation of Reliability Coefficient

The Value of r	Reading
0,80-1,00	Very High
0,60-0,79	High
0,40-0,59	Enough
0,20-0,39	Low
0,00-0,19	Very Low (not correlated)

(Sumber: Sukiman, 2012)

The results of the research data reliability analysis are presented in the table 9:

Table 9.Reliability Test

Source	S
k	15
$\Sigma p(1-P)$	2.67
Standar Deviasi	2.46
KR-20 = 0.637	

The results of the reliability analysis using the KR-20 formula, the KR-20 value is 0.637. This value is greater than 0.6. This shows that the questions submitted as an instrument are reliable or trustworthy.

1. Level of Dificulty

The level of difficulty of a test item or question (denoted by P) is the proportion of all students who answered correctly on the subject or difficulty level questions on the field examination results seen from the% percent in the Quest program output. the difficulty level category of the questions can be seen from

Table 10. Category of Problem Level of Difficulty

P Price	Question Category
0,00-0,29	Difficult
0,30-0,69	Moderate
0,70-1,00	Easy

(Sumber: Sukiman, 2012)

The results of the analysis of the difficulty level of the questions are presented in table 11.

Table 11. Test the difficulty level of the question

Number	Sum of Name	TK	Ket
1	15	0.44	Moderate
2	30	0.88	Easy
3	15	0.44	Moderate
4	17	0.50	Moderate
5	3	0.09	Difficult
6	5	0.15	Difficult
7	11	0.32	Moderate
8	26	0.76	Easy
9	30	0.88	Easy
10	29	0.85	Easy
11	30	0.88	Easy
12	10	0.29	Difficult
13	22	0.65	Moderate
14	21	0.62	Moderate
15	30	0.88	Easy

The results of the analysis of the level of difficulty obtained by the group of questions included in the category of easy difficulty level were questions 2, 8, 9, 10, 11, and 15. The group of questions that had a moderate difficulty level were questions 1, 3, 4, 7, 13, and 14. Questions that have a difficult level of difficulty are questions 5, 6, and 12.

Conclusion

Based on the results of research in the field, it can be concluded that the validity of the various tests is valid. So that students' science problem solving instruments can be used. This instrument is also suitable for use on a large scale students. The validity test also shows that the number of valid questions is 11 items and 4 items are invalid. The reliability of the question is also very good, which is equal to 0.637 which is greater than 0.6 so that it is reliable. This question instrument can be used as an alternative to one of the summative questions in school.

Acknowledgment

N/A

References

Abdullah, A. (2007). Kurikulum Pendidikan di Indonesia Sepanjang Sejarah (Suatu Tinjauan Kritis Filosofis). Jurnal Pendidikan Dan Kebudayaan. https://doi.org/10.24832/jpnk.v13i66.354

Alhunaini, S., Osman, K., & Abdurab, N. (2020). The Development and Validation of Assessment Practices of Mathematical Thinking (APMT) Instrument. Jurnal Pendidikan Malaysia, 45(02), 79–91. https://doi.org/10.17576/jpen-2020-45.02-08

Astika, I. P. W., Marhaeni, A., & Agung Parwata, I. G. L. (2019). Pengaruh Metode Pembelajaran Kwl dan Asesmen Diri terhadap Kemandirian Belajar dan Keterampilan Membaca Wacana Bahasa Bali Pada Siswa Kelas Xi Mipa Sma Negeri 1 Singaraja. Jurnal Penelitian Dan Evaluasi Pendidikan Indonesia. https://doi.org/10.23887/jpepi.v9i1.2807

Borg, W. R., & Gall, M. D. (1983). Education Research: An Introducation. Longman Inc.

Cavanagh, A. J., Aragón, O. R., Chen, X., Couch, B., Durham, M., Bobrownicki, A., Hanauer, D. I., & Graham, M. J. (2016). Student buy-in to active learning in a college science course. CBE Life Sciences Education. https://doi.org/10.1187/cbe.16-07-0212

Dharmawati, Rahayu, S., & Mahanal, S. (2016). Pengembangan Instrumen Asesmen Berpikir Kritis Untuk Siswa Smp Kelas Vii Pada Materi Interaksi Makhluk Hidup Dengan Lingkungan. Jurnal Pendidikan: Teori, Penelitian, Dan Pengembangan.

Ellianawati, E., Mufiatunnikmah, S., Setyaningsih, N. E., & Subali, B. (2020). Asesmen Multi Representasi Berbasis Keterampilan Abad Ke-21 pada Materi Gerak Lurus. Physics Education Research Journal, 2(1), 19–34.

Furnham, A. (2009). The validity of a new, self-report measure of multiple intelligence. Current Psychology, 28(4), 225–239. https://doi.org/10.1007/s12144-009-9064-z

Huijgen, T., van de Grift, W., van Boxtel, C., & Holthuis, P. (2017). Teaching historical contextualization: the construction of a reliable observation instrument. European Journal of Psychology of Education. https://doi.org/10.1007/s10212-016-0295-8

Karamustafaoğlu, S. (2011). Improving the Science Process Skills Ability of Science Student Teachers Using I Diagrams. Eurasian J. Phys. Chem. Educ.

Kusuma, M. D., Rosidin, U., Abdurrahman, A., & Suyatna, A. (2017). The Development of Higher Order Thinking Skill (Hots) Instrument Assessment In Physics Study. IOSR Journal of Research & Method in Education (IOSRJRME). https://doi.org/10.9790/7388-0701052632

Kusumah, R. G. T., Walid, A., Pitaloka, S., Dewi, P. S., & Agustriana, N. (2020). Penerapan Metode Inquiry Untuk Meningkatkan Hasil Belajar Materi Penggolongan Hewan di Kelas IV SD Seluma. JURNAL PENDIDIKAN MATEMATIKA DAN IPA, 11(1), 142–153. https://doi.org/10.26418/jpmipa.v11i1.34708

Life, W., & Instrument, Q. (2020). Kesahan dan Kebolehpercayaan Instrumen Kualiti Kehidupan Kerja Guru (Validity and Reliability of Teachers' Working Life Quality Instrument). Jurnal Pendidikan Malaysia, 45(02), 1–12. https://doi.org/10.17576/jpen-2020-45.02-01

Mourtos, N. J., Okamoto, N. D., & Rhee, J. (2004). Defining, teaching, and assessing problem solving skills. 7th UICEE Annual Conference on Engineering Education, 1–5.

Putra, B. K. B., Prayitno, B. A., & Maridi. (2018). The effectiveness of guided inquiry and instad towards students' critical thinking skills on circulatory system materials. Jurnal Pendidikan IPA Indonesia, 7(4), 476–482. https://doi.org/10.15294/jpii.v7i4.14302

Retnawati, H., Djidu, H., Kartianom, Apino, E., & Anazifa, R. D. (2018). Teachers' knowledge about higher-order thinking skills and its learning strategy. Problems of Education in the 21st Century. https://doi.org/10.33225/pec/18.76.215

Rusmansyah, Yuanita, L., Ibrahim, M., Isnawati, & Prahani, B. K. (2019). Innovative chemistry learning model: Improving the critical thinking skill and self-efficacy of pre-service chemistry teachers. Journal of Technology and Science Education, 9(1), 59–76. https://doi.org/10.3926/jotse.555

Tanujaya, B., Mumu, J., & Margono, G. (2017). The Relationship between Higher Order Thinking Skills and Academic Performance of Student in Mathematics Instruction. International Education Studies. https://doi.org/10.5539/ies.v10n11p78

Thompson, T. (2008). Mathematics teachers' interpretation of higher-order thinking in Bloom's taxonomy. International Electronic Journal of Mathematics Education.

Trisnawati, W. W., & Sari, A. K. (2019). Integrasi Keterampilan Abad 21 Dalam Modul Sociolinguistics: Keterampilan 4C (Collaboration, Communication, Critical Thinking, Dan Creativity). Jurnal Muara Pendidikan, 4(2), 455–466.

Widana, I. W. (2018). Higher Order Thinking Skills Assessment towards Critical Thinking on Mathematics Lesson. International Journal of Social Sciences and Humanities (IJSSH). https://doi.org/10.29332/ijssh.v2n1.74

Wirianto, D. (2014). Perspektif Historis Transformasi Kurikulum di Indonesia. Islamic Studies Journal.