PPSDP International Journal of Education
Volume 2 (2) (Special Issue) 05-06 July 2023, 169-178
1st PPSDP International Conference on Educational Sciences (IConEds 2023)
E-ISSN 2829-5196, P-ISSN 2830-3229

The Development Instrument of High Order Thinking Skills Assessment for the Elementary Student

Ermi Sola[1], Patta Bundu[2]
[1]Universitas Islam Negeri Alauddin Makassar, South Sulawesi, Indonesia,
[2]Universitas Negeri Makassar, South Sulawesi, Indonesia

Corresponding author e-mail: ermisola18@gmail.com

Abstract: The main objective of this research is to provide High Order Thinking Skills (HOTS) instrument assessment at elementary school. The development model used some phases: develop blueprint test, write the item test, verify the item test, tryout 1, analyses/ revision, tryout 2, analyze test result, and finalize the instrument. Analysis technique for content and construct validity using index Grogery formula at ≥ 0,75 index for relevancy (validity). For empirical validity Karl-Pearson correlation formula, and reliability using coefficient α (Cronbach Alpha). The research result shows that HOTS instrument assessment is valid and reliable, so that it can be used in teaching – learning activities to assess HOTS of the students at elementary school.

Keywords: Assessment, Elementary Student, High Order Thinking Skill

A. Introduction

A teacher must have pedagogic competence to manage education to achieve learner-centered learning. One indicator is learning with good assessment, feedback, and reporting. Assessment must be an integrated part of the learning process, facilitating learning, and providing holistic information as feedback for educators, students, and parents, in order to guide them in determining further learning strategies. The conditions for a good assessment must include learning processes and products so that complaints that have arisen that assessments only measure a low cognitive level can be overcome. It is suspected that so far most elementary school teachers tend to still measure lower order thinking skills (LOTS) which are at cognitive level C1 (memory) and C2 (knowledge). A higher level in questions of higher order thinking skills (HOTS) is still rarely done. made out of context. The questions prepared by the teacher usually measure recall skills. When viewed from the context, most of them use contexts in the classroom and are very theoretical, and rarely use contexts outside the classroom. So that it does not reveal the relationship between knowledge obtained in learning and real situations in everyday life (not contextual).

A survey conducted at several elementary schools in the city of Makassar showed that almost all of them had not implemented the HOTS assessment in learning. The teachers already understand that the questions will be much more meaningful if the content starts from C3 and above. The Ministry of National Education (2014) has published a HOTS assessment guide but its implementation has not yet been seen in the field. One of the serious problems is that teachers have not been able to develop HOTS-oriented learning assessments. It is at this level that the development of HOTS instruments is very important in learning in elementary schools.

The form of questions that students often work on also seems to be dominated by memory in Low Order Thinking Skills (LOTS). There is a marked lack of the drive for critical and creative thinking required in HOTS. The following is a comparison between questions in grade IV elementary school taken from science books and questions developed in the Elementary Science Curriculum Vancouver Canada. The subject matter is the same but the form of assessment is different.

From Book Class IV Elementary School
"Insects like grasshoppers have ..... pairs of legs"
A. 2 pairs C. 3 pairs
B. 4 pairs D. 5 pairs

From Elementary Science Curriculum – Grade 4. "Insects have 3 pairs of legs. Which of the following pictures is not an insect?"

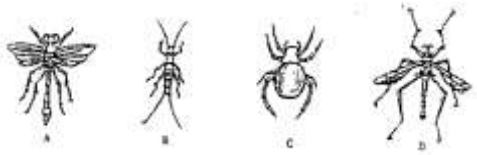

Figure 1. **Insects**

In addition, the results of the international study Program for International Student Assessment (PISA) as quoted by Saraswati et al. (2021) show reading literacy, mathematical literacy, and scientific literacy achievements achieved by Indonesian students. very low. In general, the ability of Indonesian students is very low in: (1) understanding complex information; (2) theory, analysis, and problem solving; (3) use of tools, procedures and problem solving; and (4) conduct investigations.

Even more concerning is that Elementary School Teacher Education (PGSD) has not shown the ability for students to solve HIOTS questions, which in the future they will become teachers in elementary schools. Research by Irfan et al. (2022) shows that the ability to solve HOTS questions for PGSD FIP Makassar State University students is in the following categories:

Table 1. Ability to Solve HOTS Questions for PGSD Students Category Achievement Courses

Courses	Achievement	Category
Science Education	66.67%	Moderate
Indonesian Language	57.78%	Low
Mathematics	46.74%	Low

Source: Research Results of Irfan et al. (2022)

The ability to solve HOTS questions at the elementary school level is even lower, as Irfan et al. (2022) found that the ability to solve HOTS questions is less than 40% of fifth grade elementary school students who do well. The results of interviews with several elementary school teachers at the KKG indicated that they had not been able to develop the HOTS assessment instrument because not only was there a lack of training for this, but there was also a lack of examples or guidelines for preparing HOTS questions in learning.

In general, the development of test instruments is carried out in order to produce valid and quality instruments. There are many variations of the instrument development steps but the content is much the same. Mardapi (2012) suggests 9 steps namely: (a) compiling test specifications, (b) writing test questions, (c) reviewing test questions, (d) conducting test trials, (e) analyzing test items, (f) fixing the test, (g) assembling the test, (h) carrying out the test, and (i) interpreting the test results.

Meanwhile Lehmann & Mehrens (1987) suggest to follow when developing an instrument 1) Set learning goals carefully; 2) Prepare a specification table, as a reference when writing instrument items; 3) Formulate items that are clear, unambiguous, and unambiguous; 4) Make sure the items are from the material that has been taught; 5) Try to be free from bias, because of tradition, gender, and so on; 6) Prepare the scoring key when writing the item as well, if it's for study results; 7) Prepare more items than needed; and 8) Prepare these items as early as possible, so there is time to revise.

As for the things that must be considered in developing the test according to Arifin (2012.), namely: aspects to be measured, the compiler, the purpose of using the test, samples, validity and reliability, administration, how to score, answer keys, raw score tables, and interpretation. With regard to HOTS, Brookhart (2010) suggests that to define Higher Order Thinking Skills (HOTS) it must be aligned with the standard content and learning objectives in the classroom. Brookhart (2010) classifies HOTS into three categories: (1) those who define HOTS in the sense of transfer, (2) those who define it as critical thinking, and (3) those who define it as problem solving. The transfer category is indicated as "the meaningful learning) that requires students not

only to remember but also to make sense of and be able to use what they have learned (Anderson & Krathwohl, 2001). Critical thinking is defined as "reasonable, reflective thinking that is focused on deciding what to believe or do. Reasoning, questioning and investigating, observing and describing, comparing and connecting, finding complexity, and exploring viewpoints (Norris, 1989). While the category of problem solving is defined as a thought process in solving problems so that goals are achieved (Nitko & Brookhart, 2007).

In Ministry of Education and Culture Number 21 of 2016 concerning Content Standards for Elementary and Secondary Education it is explicitly stated that the learning outcomes in the knowledge domain follow Bloom's Taxonomy which has been revised by Anderson and Krathwohl (2001) consisting of the ability to: know (C1), understand (understand ing-C2), applying (aplying-C3), analyzing (analyzing-C4), evaluating (evaluating-C5), and creating (creating-C6). HOTS questions measure abilities in the realm of analyzing (analyzing-C4), evaluating (evaluating-C5), and creating (creating-C6). Furthermore, Karthwhol and Anderson (2001) added dimensions of knowledge, starting at the factual, conceptual, procedural, and metacognitive levels.

If these two concepts, namely the cognitive level and the knowledge dimension, are combined, the HOTS dimension will be at levels C4 to C6 and on the conceptual, procedural, or metacognition dimensions. The combination can be seen in more detail, as follows:

Table 2. Dimension HOT

Knowledge	**Metacognition**	LOTS AREA					
	Procedural						
	Conceptual						
	Factual						
	Cognitive	**C1**	**C2**	**C3**	**C4**	**C5**	**C6**

According to Brookhart (2010) there are at least three basic principles for carrying out HOTS assessments in learning: (1) using introductory material or allowing access to source material, (2) using new material, and (3) paying attention to cognitive complexity and difficulties separately. That is, in order for the HOTS assessment to be carried out, (1) use introductory material or provide opportunities to access material, (2) use relatively new material, and (3) make perfection of cognitive complexity and difficulties.

Based on a brief description of the assessment conditions that occurred in the field, the problem formulation is structured as follows:

1. What are the steps for developing HOTS assessment instruments for elementary school students?
2. Does the HOTS assessment instrument that is produced meet the valid and reliable criteria in learning in elementary schools?

B. Methods

This type of research adopted a development research model, which produced a HOTS-oriented learning assessment instrument for elementary school students. The stages of instrument development follow the development steps as shown in the following diagram.

Table 3. Steps of D

Development Stages	Main Activity
Develop test specifications	• define test objectives, • arrange the grid and choose the kinds of the test
Writing test questions	Writing test items in the form of a description of the indicators according to the details on the grid.
Reviewing test questions	• Proof reading • Expert test (content and construct)
Test trials (1)	• Small scale trials (1 SD)
Analyze the test items. (1)	• Validity • Reliability • Revision of repairable items,
Revision/Reassembly	• Reassemble the instrument (order number, layout)
Test trials (2)	• Wide scale trials (3 SD
Item Analysis (2)	• Validity • Reliability
Assembling the Final Instrument	• Calibrated instrument

Content and construct validity were validated by two experts in their fields, by calculating based on the validity index from Gregory (Azwar, 2017). Index ≥ 0.75 is declared relevant (valid) provided that there are no items from the assessment aspect that get a score of 1, and the difference in the assessment of each item is a maximum of 2 points. The formula used is as follows.

Table 4. Gregory's Validator Relevance

		Validator 1	
		Low (1,2)	High (3,4)
	Low (1,2)	**A**	**B**
Validator 2	High (3,4)	**C**	**D**

Content validity coefficient can be calculated using the formula:

$$\text{Valliditas Isi} = \frac{D}{(A + B + C + D)}$$

Emperical validity was analyzed using Product Moment correlation by Karl Pearson, as follows:

$$r = \frac{N\Sigma XY - (\Sigma X)(\Sigma Y)}{\sqrt{\{N\Sigma X^2 - (\Sigma X)^2\}\{N\Sigma Y^2 - (\Sigma Y)^2\}}}$$

r = coorelation coeficient
ΣXY = sum of score X time Y
ΣX = sum of X score
ΣY = sum of Y score
ΣX^2 = sum of square X score
ΣY^2 = sum of square Y score
N = number of students

Reliability of instrument was ditermined using internal consistency of *Cronbach's Alpha* (α. coeficient)

$$\alpha = \frac{k}{k-1}\left(1 - \frac{\Sigma\sigma_i^2}{\sigma_x^2}\right)$$

Nore: k = numbers of items. σ_i^2 = variant items test. σ_x^2 = variant of total score

C. Results and Discussion

Based on the stages of instrument development used, the research results are grouped into two stages, namely the stage of instrument development and the results of testing the HOTS nuanced assessment instrument.

Development Stages

At the development stage, a HOTS assessment instrument was produced which was not based on the contents of each basic competency in the curriculum. There are two fundamental considerations, namely (1) enabling teachers to later make examples of HOTS assessments for each subject, and (2) focusing more on the HOTS assessment criteria themselves. There are three HOTS assessment criteria used, namely: cognitive level and knowledge level, HOTS indicators, and HOTS question stimulation. Here are some examples made in question cards provide logical and scientific findings of the study.

Question Card No. 1

Question: Look at the picture below: (1) who are they in tis picture? (2) what is happening? (3) what will they do next?

Scoring Rubric/Guidelines
Logical answers according to the image interpretation. Score = 2
The less logical answer still interprets the image. Score = 1.
Illogical/empty answer Score = 0.
Maximum score 3 x 2 = 6

Question Card No. 2

Question: Fill in the 10 boxes below with one word. Make three different sentences, using these words. You may use these words repeatedly, but may not use words other than those in the box.

No.	Word	No.	Word
1		6	
2		7	
3		8	
4		9	
5		10	

a. …….....
b. …….....
c. …….....

Rubric
Sentences made correctly based on the letters in the box, score = 2
Sentences that are made incorrectly, = score 1
Improper/blank sentences, score = 0
Maximum score = 6.

Prior to field trials, the instrument was validated in terms of content and constructs by two expert validators. The results were analyzed using the Grogery format (standard coefficient 0.75). The result obtained = 0.86 which indicates that the HOTS Assessment Instrument meets the valid criteria in terms of content and construct (0.86 > 0.75).

Instrument Test Results

The instrument was tested on class V elementary school with 30 students. The validity of the instrument is based on the Pearson Product Moment Correlation

formula. Instrument items are declared valid if the Pearson correlation coefficient ≥ 0.30. Data were analyzed manually with the help of a simple Excel program. The full results are as follows.

Table 5. Data Analyzed

Number of Items	Coefficient Correlation	Criteria	Number of Items	Coefficient Correlation	Criteria
1.	0.764166	Valid	6.	0.296393	Valid
2.	0.427002	Valid	7.	0.627426	Valid
3.	0.410573	Valid	8.	0.529159	Valid
4.	0.665969	Valid	9.	0.719259	Valid
5.	0.524648	Valid	10.	0.569375	Valid

The reliability of the instrument using the Alpha Cronbach formula and the HOTS Assessment Instrument was stated to be reliable with r = 0.70

Discussion

The development of the instrument needs to be identified because even though there are similarities, there are still differences in the stages between one expert and another. In this study using stages similar to the pattern from Mardapi (2012) and Arifin (2012). As described in the Development Stages research method through compiling test specifications, writing test items/items, reviewing test items, testing, analysis/revision, final instruments.

The advantage of this stage is that the activities at each step of its development are very clear. Even in its development it can accommodate the HOTS indicators put forward by the Brookhart's Instrument (2010) where there are at least three basic principles for carrying out HOTS assessments in learning: (1) use introductory material or allow access to resource material, (2) use novel material, and (3) attend separately to cognitive complexity and difficulty.

The advantages of the resulting HOTS instrument can be used as a guide for teachers to implement the HOTS assessment instrument in classroom learning. The obstacles that arose when assembling the initial test draft were due to the large number of HOTS indicators that had to be accommodated. Sani R.A (2019) focuses on the HOTS dimension of creativity which contains 4 main pillars such as: fluency, flexibility, elaboration, and originality.

As a consequence, there are several HOTS dimensions that were not accommodated in this study but were more focused on indicators that accommodated the HOTS assessment itself such as cognitive level and knowledge dimensions. This is also

based on the opinion of Sani, R.A. (2019) suggested six core aspects of HOTS, namely disposition, criteria, arguments, reasoning, point of view, and application procedures. Next, each aspect is described as follows:

Table 6. Core Aspects of HOTS

Core Aspect	Description
Disposition	Open-minded, value evidence/reasoning
Criteria	Conditions that the statement must meet
Arguments	Logical evidence to support a statement
Reasoning	There is a logical relationship between statements and data
View	Seeing phenomena from various points of view
Procedure	Analyze the thought process for one case

The HOTS instruments produced have been tested for validity and reliability, although the format used for each instrument is different. However, it still needs to be tested on a wider scale and if possible modified according to the basic competencies of each field of study for which the HOTS assessment instrument will be made.

D. Conclusion

The development of learning assessment instruments produces a HOTS Learning Assessment model that fulfills the main dimensions of creativity, critical thinking, and problem solving. The assessment model meets a high level of validity and reliability for use in learning in elementary schools. It is suggested that in the future instruments for HOTS dimensions will be developed separately such as critical thinking, creativity, problem solving, and decision making.

E. Acknowledgments

Thank you to the Chairperson of the Research and Community Service Institute, Makassar State University (UNM) who provided funding for this research, the Director of the UNM Postgraduate Program for granting research permits, and the Principal of Minasa Upa State Elementary School for preparing his school for the HOTS assessment instrument trials.

References

Anderson, L. W., & Krathwohl, D. R. (2001). *A Taxonomy for Learning, Teaching, and Assessing, A Revision of Bloom's Taxonomy of Educational Objectives*. New York: Addison Wesley Longman Inc.

Arifin, Z. (2012). *Evaluation of Learning Principles, Techniques, Procedures*. Bandung: Remaja Rosdakarya.

Azwar, S. 2017. *Preparation of Psychological Scales (2nd ed., print XI)*. Yogyakarta: Pustaka Pelajar.

Brookhart, S. M. (2010). *How to Assess Higher-order Thinking Skills in Your Classroom*. Ascd. Virginia: ASDS' Online Store.

Irfan, M., Patta, R., Rahman, A., Bundu, P., & Amran, M. (2022). PKM Training on Preparing Learning Assessments Oriented to Higher Order Thinking Skills (HOTS). *CARADDE: Jurnal Pengabdian Kepada Masyarakat*, *5*(2), 301–310.

Lehmann, I. J., & Mehrens, W. A. (1987). Using Teacher-Made Measurement Devices. *NASSP Bulletin*, *71*(496), 36–44. https://doi.org/10.1177/019263658707149605

Mardapi, D. 2012. *Educational Assessment & Evaluation Measurement*. Yogyakarta: Nuha Litera

Nitko, J. A. & Brookhart, S. M. 2007. *Educational Assessment of Student*. Fifth edition. New Jersey: Merrill Prentice Hall.

Norris, S. 1989. Can We Test Validity for Critical Thinking. *Educational Researcher*, 18 (9), pp.21 – 26.

Sani, R.A. 2019. *Learning based on HOTS (Higher Order Thinking Skills)*. Tangerang: Tira Smart

Saraswati, Y., Sifak Indana, & Elok Sudibyo. (2021). Science Literacy Profile of Junior High School Students Based on Knowledge, Competence, Cognitive, and Context Aspects. *IJORER : International Journal of Recent Educational Research*, *2*(3), 329–341. https://doi.org/10.46245/ijorer.v2i3.118