

Higher order thinking skills-based reading literacy assessment instrument: An Indonesian context

Vismaia Sabariah Damaianti*, Yunus Abidin, and Rosita Rahma

Department of Indonesian Language Education, Faculty of Language and Literature Education,
Universitas Pendidikan Indonesia, Jl. Dr. Setiabudhi 229, Bandung, West Java, Indonesia

ABSTRACT

Learning theories have seen the development of students' higher order thinking skills a quintessential educational goal for all students, as the absence of such skills in learning leads to students' difficulty in answering questions that are analytical, critical, creative, and problem-solving. What is more, the prevailing literacy scoring instruments have yet to take into account the Indonesian cultural context despite the fact that culture is such an important vehicle in strengthening the identity of a nation. In order to address this problem, employing a research-and-development method, a HOTS-based reading literacy scoring device model was proposed in this research. In the development stage, the model was tested to 476 junior high school students in two separate islands in Indonesia: Java and Bali. The results of the qualitative assessment from the experts showed that the product developed in principle had fulfilled the requirements. Meanwhile, the validity and reliability test results demonstrated that the instrument under investigation had met the requirements as a standardized reading literacy assessment product. Implicationally, therefore, the proposed model can be utilized in assessing students' reading skill in Indonesian contexts.

Keywords: HOTS; Indonesian cultural context; reading literacy assessment instrument

First Received:

1 April 2020

Revised:

19 July 2020

Accepted:

4 August 2020

Final Proof Received:

25 September 2020

Published:

30 September 2020

How to cite (in APA style):

Damaianti, V. S., Abidin, Y., & Rahma, R. (2020). Higher order thinking skills-based reading literacy assessment instrument: An Indonesian context. *Indonesian Journal of Applied Linguistics*, 10(2), 513-525. <https://doi.org/10.17509/ijal.v10i2.28600>

INTRODUCTION

Literacy plays a vital role in people's life as one who knows how to read, write and count will contribute more significantly to society and can understand the world (Barton et al., 2000; Gunes, 1997). The concept of literacy has developed from a simple concept as the ability to read and write to the ability to apply various competencies and skills in life. Literacy skill is crucial to keep abreast with technological and sociocultural developments in the 21st century. Of an empirical interest is how to integrate literacy training in all learning materials to foster student capacity (Ng & Graham, 2017; Greenleaf et al., 2010; Morocco et al., 2008).

Many obstacles are abound in measures to develop literacy skills, especially reading literacy.

Currently, the reading literacy level of junior high school students in Indonesia sits in the low category (OECD, 2016; TIMSS & PIRLS, 2012). Thus, the need exists for a new design in a comprehensive literacy learning system by considering the quality of all learning components.

One component that influences the quality of the literacy learning system is the learning assessment standard. Current reading literacy education policies should prioritize standardized assessments, so that control over many factors can influence outcomes and learning processes (Davies & Bansel, 2007; Edglossary, 2014). As an initial step, efforts should be made to develop student literacy competencies by developing learning standards that are relevant to the demands of time.

* Corresponding Author
Email: vismaia@upi.edu

Learning standards can be obtained through an appropriate assessment system. Any initiative to develop good learning must begin with the development of an appropriate assessment system, so that assessment can be used as a guide to the learning process (Fulcher & Davidson, 2007; Picone-Zocchia, 2009; Trilling & Fadel, 2009; Weeden et al., 2003; Wormeli, 2018). In order to have meaning and benefit, all literacy assessments must provide some added value for teaching and learning (International Reading Association & National Council of Teachers of English [IRA & NCTE], 2010). One rating system that can be used is higher order thinking skills (HOTS). The implementation of HOTS will have an impact on improving students' thinking skills and performance, thereby helping students improve understanding of content in the text (Brookhart, 2010). Unfortunately, schools in Indonesia have not yet implemented the HOTS assessment which refers to the development of students' critical and creative thinking skills (Abidin, 2013).

Reading literacy assessment must be adjusted to the learning objectives and diversity of student backgrounds. The recent development of literacy studies indicates that the objectives of literacy education must pertain to sociocultural theories and cultural practices. From a sociocultural perspective, it is evident that in assessment activities, an individual strives to understand meaning by bringing his cultural background (Barton et al., 2000; Martin-Jones & Jones, 2000; Willis et al., 2013). Thus, literacy assessment should be supported by strengthening the context of cultural settings.

Among a number of research studies related to the development of standard assessments in learning to read was conducted by Murray et al. (2011). Their research demonstrates that assessments that are considered capable of improving students' reading comprehension are manifested into five types of assessment including (1) letter-sound correspondences, (2) word recognition, (3) decoding, (4) fluency, and (5) comprehension. Another study was performed by Alonzo et al. (2009). The resulting reading assessment model can be categorized as a standard assessment with several levels of variation including literal level, inferential level, and evaluation level. This research suggests that the assessment model can identify skills that are difficult for students to master. The results can determine more appropriate learning goals for reading. Smagorinsky (2009) conducted a study to develop a standard reading assessment model based on reading culture. He proposed three types of reading standard assessments in three dimensions, namely the Self-Evident Construct dimension, the Discrete Act dimension, and the Cultural Act dimension. His research shows reading assessments should be developed by considering children's

ability to construct knowledge based on their own experiences. In a similar vein, Provost et al. (2009) developed a reading assessment model based on Informal Reading Inventories to measure the ability to read in several stages, namely (1) measuring comprehension, (2) calculating comprehension, (3) error analysis, and (4) determining understanding determination.

Although a variety of reading assessment models have been proposed in the aforementioned previous studies, no studies have specifically developed a higher order thinking skills-based reading literacy assessment. In light of that, this research focused on developing a HOTS-based reading literacy assessment tool in the Indonesian cultural context in an effort to develop students' literacy skills, especially reading literacy.

Reading literacy assessment

Reading literacy assessment is defined as a way of assessing what students know and do from their reading activities, how to interpret assessment results, how to apply assessment results, and how to improve learning based on assessment results. The more teachers know about literacy assessment, the more progress can be made to make decisions designed to improve students' future (Crandall et al., 2016; Webb, 2002). Thus, literacy assessment includes a series of procedures to help teachers make learning decisions. Efforts to assess the ability to read literacy must be done by using an appropriate reading literacy assessment instrument. In connection with this, PISA literacy problems can be used as a reference in developing standardized literacy measurement tools. The PISA International student assessment program is one of the largest international scale efforts that has been launched to assess students' scientific literacy. Such international assessments will have a major impact on the science education policies of participating countries (Lau, 2009). In line with this, efforts to develop reading literacy assessment instruments can be made by referring to concepts, frameworks, and examples of PISA questions. PISA questions designed to measure literacy can be divided into three main aspects. The first aspect, namely the situation refers to various contexts or objectives. The second aspect, namely the text, refers to diverse reading material. The third aspect refers to a cognitive approach that determines how the reader engages with the text. In PISA, features of text variables and aspects (but not from situation variables) are also manipulated (OECD, 2016).

Based on the test structure developed by PISA, reading questions measure more reasoning, problem solving, argumentation, and communication skills than questions that measure memory and comprehension abilities. Furthermore, PISA questions also measure the level of students' ability to solve problems that require higher reasoning or

HOT skills. In the 1980s many experts considered the importance of standardized assessments with higher-level thinking skills indicators. The discovery of the right solution to complex problems is obtained through a higher-level thinking process. Naturally, teaching high-level thinking can help students to become skilled students in their lives and help students improve their understanding of content with high-level thinking (DeVries & Kohlberg, 1987; McDavitt, 1994; Son & VanSickle, 1993).

The concept of reading literacy

In the concept of literacy, reading is interpreted as an effort to understand, use, reflect, and involve various types of texts in order to achieve a goal that is to develop one's knowledge and potential and to participate in society. The focus is that literacy reading is how individuals make meaning through interaction with text, the process of reading involves a sociocultural context (Frankel et al. 2016; Purcell-Gates et al., 2016). Based on this definition, reading is interpreted as an activity of building meaning, using information from reading directly in life, and linking information from the text with the experience of the reader in life, and linking information from the text with the experience of the reader (Frankel et al., 2016; Snow, 2002). Reading in this sense really requires the ability to analyze and synthesize information so that the resulting understanding has a complex structure of meaning. The definition of reading must go further by paying attention to processes as they occur in context. This extended definition provides a perspective that requires a shift in focus from reading to literacy (Frankel et al., 2016; Purcell-Gates et al., 2016). From this perspective, there are differences between "learning to read" and "reading to learn". Reading literacy is the ability to read to learn, which is a set of skills that equip readers to deal with problems in accordance with text understanding and context becomes increasingly problematic because teaching reading as a set of general skills and strategies does not equip readers to deal with text and context demands (Pearson & Cervetti, 2013).

In line with PISA's view, reading ability is more related to the concept of careful reading. Reading carefully at the beginning of its appearance is said to be the technical analysis of texts. In line with this conception, careful reading emphasizes more on strategies to understand how the writer presents his ideas, pay attention to the choice of words made by the author, and understand the messages that are converted in important features contained in the discourse. In informational and argumentative texts, the reader also needs to test the author's statement and the evidence the author uses to strengthen his statement. Sisson and Sisson (2014) state that careful reading is a process of reading that is carried out repeatedly on complex texts that aim to achieve three stages of

understanding namely literal understanding, inferential understanding, and evaluative understanding. Lapp et al. (2015) remark that careful reading is a very important reading process because it is in line with today's literacy learning standards. Through careful reading activities, readers are expected to develop their abilities in (1) understanding the general contents of the text in general; (2) finding the key details of the text; (3) developing vocabulary and the structure of texts; (4) understanding the writer's purpose; (5) drawing inferences of reading content; and (6) developing opinions, arguments, and connecting various texts. Based on this careful reading function, the purpose of reading is not only to gain a superficial understanding of complex texts but also to evaluate a variety of complex texts.

The concept of careful reading was also put forward by Benjamin and Hugelmeier (2013) that contend that careful reading is a short, complex text reading activity undertaken to find a proof contained in a text. The evidence contained in the text can be presented either directly or indirectly. Based on this understanding, careful reading is to arrive at a deep understanding that is accompanied by real evidence contained in the text. That careful reading is a reading activity to gain a deep understanding of a text. Tantillo (2012) defines reading more precisely as a systematic practical activity in analyzing texts to gain a deep understanding. Based on the above definition, reading literacy is an activity that emphasizes the acquisition of a deep understanding of something involving high-level thinking skills. Thus, reading literacy is not just understanding a reading text but also synthesizing reading texts even further the ability to use information and evaluate information. Therefore, reading literacy is an ability that must continue to be developed throughout students' academic life.

HOTS-Based literacy reading assessment

The need to set higher-order thinking skills standards has been documented throughout the 1980s and 1990s. In fact, Anderson (1985) reports that the Reading Commission called *Becoming a Nation of Readers* makes educational excellence through assessment with high-level thinking standards. Florida Department of Education (1996-1997) states learning goals that are based on higher-order thinking enable students to make wise and healthy life decisions. Likewise, Secretary's Commission on Achieving Necessary Skills (SCANS) (1991) argues that education is said to be successful if it produces students who can think creatively, make decisions, solve problems, visualize, know how and reasons for learning (SCANS, 1991). There are several standard indicators of reading literacy ability that can be used as a reference in making HOTS-based reading literacy measurement tools. Among them are critical

thinking abilities, creative thinking abilities, metacognitive abilities, procedural thinking skills, schematic abilities, and the ability to understand visual images.

One has the ability to think critically if one is able to provide an assessment of various solutions to problems (Crowl et al., 1997; Lewis & Smith, 1993). By thinking critically, a reader can think reflectively and make sense in evaluating evidence from an argumentative statement (Crowl et al., 1997; Facione, 1998; Lewis & Smith, 1993; Patrick, 1986). When one thinks of solutions to problems, one needs a creative process. Creativity is the ability to produce new ideas. Someone who has creativity can use basic concepts or rules in new contexts and situations. In overcoming problems, one who thinks creatively is able to involve relevant concepts and then integrate new information into the concept (Crowl et al., 1997; Sternberg & Davidson, 1995). A problem is a situation when one wants to get what one wants but does not know what action to take. The problem solving is the success in getting various decisions (Crowl et al., 1997). The level of thinking ability also depends on how one responds to contexts in the real world that challenges the thought process. One's success in thinking at a high level depends on one's ability to apply, develop, and update knowledge according to contexts and situations.

Another variable which is an indicator of the ability to think at a higher level is the ability of metacognition. Metacognition is the ability to monitor and recognize oneself through the thought process. With the ability to think at a high level, one can correct oneself as the impact of one's understanding of reading. Even with metacognitive abilities, one will have confidence that one is able to exceed the abilities of other individuals (Crowl et al., 1997). Furthermore, indicators of higher order thinking ability are part of procedural thinking. The application of procedural knowledge which also involves analysis and synthesis can be considered

high-level thinking skills (Huot, 1995). Making links, developing maps, and compiling the grid are some of the capabilities of procedural understanding. In interpreting meaning, when reading one also uses the ability to think at a higher level through the merging of information from the text with the schemata one already has, and the ability to think at a higher level is related to the ability to understand the text of visual images.

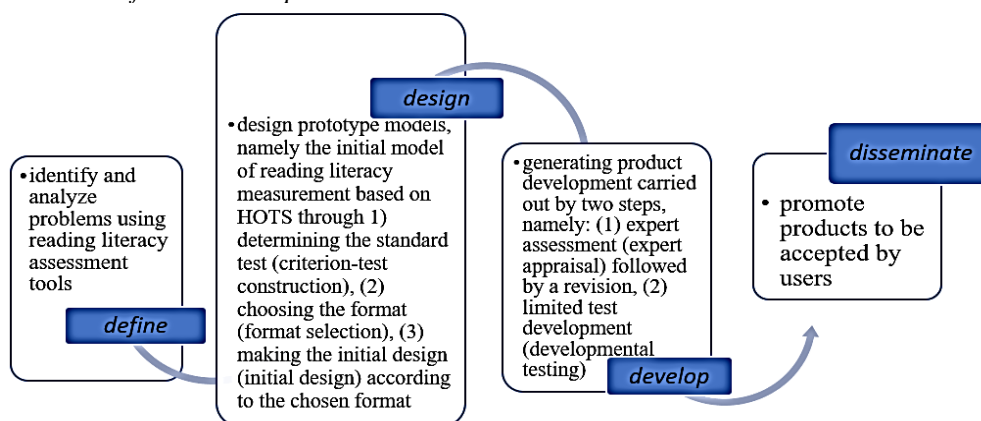
Cultural contexts in reading literacy

It is important to note that reading literacy is developed in a cultural context, reading literacy learning is learning to read words and cultural signs (Snow, 2002). Therefore, reading literacy assessments should also be adjusted and linked to cultural settings (Cole, 1998; McQueen, & Mendelovits, 2003; Vygotsky, 1978). Cultural elements relating to the setting and context of Indonesian life include elements of language, knowledge systems, social systems or social organizations, living equipment systems and technology, livelihood systems, religious systems, and art systems (Koentjaraningrat, 1988). The seven cultures can be classified as material and non-material cultures (Barkan, 2011). Nonmaterial cultures such as language, knowledge systems, social systems, and religious systems. The material cultural element includes all the physical objects of society, such as the system of living equipment and technology.

METHOD

This research was carried out based on the Research and Development step through the 4-D model, namely the steps to define, design, develop, and disseminate (Trianto, 2011). This model was chosen because the concept is in harmony with the steps of developing learning tools, including learning measurement tools as products produced in this research. There are four steps described in Figure 1.

Figure 1
Flowchart of Research Steps



Research location

This research was undertaken in various regions of Indonesia. To facilitate the national development process, two operational research areas were established, namely Java and outside Java. For development studies, six schools were chosen in the provinces of West Java, East Java and Bali.

Research subject

In connection with the research step, the data collected came from the results of expert validation (expert appraisal) and from the results of the test implementation. The experts in question were literacy experts and learning experts. There were 476 high school students from the following schools as can be seen in Table 1.

Table 1

Research Subject

No.	Schools	Students
1.	SMPN X Bandung	118
2.	SMP Y Bandung	72
3.	SMPN Z Bandung	86
4.	SMPN R Malang	80
5.	SMPIT S Malang	56
6.	SMP T Bali	64
	Total	476

Research instruments

There were four instruments utilized to gather the research data. At the stage of defining the instruments, the instruments were (1) a questionnaire and (2) interview guidelines to collect data from teachers about the problem of using literacy measurement tools in schools. Meanwhile, at the developing stage, the instruments used were (3) expert appraisal and (4) HOTS-based reading literacy measurement tools for the developmental testing phase. This expert assessment instrument was aimed at getting an overview of the accuracy of the measuring instruments developed in this study. The appraisal grid arranged can be seen in Table 2.

Table 2

Item Description of the HOTS-Based Reading Ability Literacy Assessment Instrument in the Indonesian Context of Junior High School

Subcomponent	Matter
Material Accuracy (Discourse, Fig. Illustration)	<ol style="list-style-type: none"> 1. Selection of discourse, pictures, and illustrations is in accordance with the competencies of students. 2. Selection of discourse, pictures, and illustrations is appropriate for fulfilling the HOTS dimension assessment. 3. Choice of discourse, pictures and illustrations is appropriate for the fulfillment of the Indonesian cultural context.
Accuracy of Material (questions / stem and answers)	<ol style="list-style-type: none"> 1. Questions are in accordance with student competencies. 2. Questions support the assessment based on the HOTS dimension that must be achieved. 3. Questions support understanding Indonesian cultural context, life skills, and future perspectives.
Presentation Techniques (Discourse, Pictures and Illustrations)	<ol style="list-style-type: none"> 1. Use of sentences in discourse meets the requirements of effectiveness and efficiency. 2. Presentation of the contents of the discourse meets the spelling rules requirements set forth in the General Guidelines for Indonesian Spelling. 3. Pictures and illustrations are easy to understand. 4. Images and illustrations are attractively presented.
Presentation Techniques (Questions / Stems and Answers)	<ol style="list-style-type: none"> 1. Questions and answers meet the requirements of good question writing techniques. 2. Questions fulfill the balance requirements based on the distribution of indicators. 3. Answers fulfill the balance requirements of the answer key. 4. There is adequate distraction.
Supporting Assessment Materials	<ol style="list-style-type: none"> 1. The materials conform to the development of science 2. The materials do not contain elements of pornography, extremism, radicalism, violence, racial intolerance, gender bias, plagiarism, and other deviations

Furthermore, the most important instrument is the HOTS-based Literacy Reading Instrument with a Cultural Context. In this measuring instrument, there are several components as the constructor,

namely the higher order component thinking skills, Indonesian culture, and types of text. Table 3 presents the HOTS-based reading literacy instrument product developed.

Table 3
HOTS-Based Reading Literacy Assessment Instrument in the Indonesian Context

HOTS indicator	Indonesian-ness	Text type
Cognitive (knowledge)	knowledge system	argumentation
Comprehension understanding, drawing conclusions, generalizing, (including understanding insight)	social system / social organization (social values, tolerance, mutual cooperation, cooperation, kinship, justice)	narrative
Critical thinking including rational thinking	living equipment systems and technology	fantasy story
Creative thinking Novelty, elaboration, smoothness, solving unusual problems, evaluating and finding new ways to solve problems	living livelihood system	exposition
Scaffolding	religious system and values	advertisement
Schemata	Art	observation report
Inquiry or Discovery	alignment value	short story
Metacognition self-regulation, ability, adaptability, strategy, getting things done, thinking management, memorizing time, in a short, mind map	the value of territorial integrity	exposition
Scripts Procedures that require reasoning. Novel analysis. Literary case (prose poem)	the value of national unity	
Graphic frame	value of independence	
Transfer	Ability to apply thinking skills taught separately to any subject	Procedure

FINDINGS AND DISCUSSION

This research was conducted to produce a HOTS-based reading literacy measurement tool appropriate within the Indonesian cultural context. The product creation process was carried out through four stages,

namely defining, designing, developing, and disseminating stages. The following is an example of reading literacy assessment instrument (see Figure 2) with HOTS-based being developed along with the development process mechanism.

Question no. 4

Please read the following text carefully!



<https://www.muttaqin.id/2017/09/soal-dan-jawaban-pelestarian-hewan-tumbuhan>

As an endangered species, the “Cendrawasih” bird is famous for its beauty and has a variety of colors. This is due to rampant hunting without services with maximum care in the maintenance process. This bird originating from Papua is also rarely found and almost extinct, especially due to the amount of damage to habitats due to natural changes, disasters, and man-made. However, as a generation that cares about the

beauty of nature and flora and fauna in Indonesia, it is not too late for us to conserve them. The slightest action taken will certainly be beneficial for its sustainability. Source: <https://arenahewan.com/>

Question:

Why do we need to maintain the existence and survival of the “Cendrawasih” bird?

- A. The “Cendrawasih” bird is inherited from our previous ancestors.
- B. Cendrawasih birds are protected animals.
- C. Cendrawasih birds must continue to be cared for and conserved.
- D. The “Cendrawasih” bird has suffered from habitat damage.
- E. Cendrawasih birds are rarely found and threatened with extinction.

Figure 2

Reading Literacy Assessment Instrument with HOTS-Based Development Process Flow

<div style="border: 1px solid black; padding: 5px; width: fit-content; margin: 10px auto;">Question number</div>	<p>The questions are sorted from easy to difficult.</p>
<div style="border: 1px solid black; padding: 5px; width: fit-content; margin: 10px auto;">Question instruction</div>	<p>The instructions provide stimulation to the testee on how to read the text properly.</p>
<p style="text-align: center;">Question Case</p> <div style="border: 1px solid black; padding: 10px; margin: 10px auto; width: 80%;"> <div style="border: 1px solid black; padding: 5px; margin-bottom: 10px; text-align: center;">Picture</div> <div style="border: 1px solid black; padding: 5px; text-align: center;">Text</div> </div>	<p>This section is a problem case. Each question begins with a case in the form of text that students must read.</p> <ul style="list-style-type: none"> a. Pictures are presented at the beginning. It aims to build the context of the reader before reading the text. b. Images are selected in the form of photographs of certain objects to provide an overview of factual data. c. Images are displayed in attractive colors so that the testees get a sense of pleasure. <hr/> <ul style="list-style-type: none"> a. The type of text is chosen in a variety of ways, namely argumentation, exposition, narrative, fantasy, short story, procedure, observation, and advertisement. b. Each text contains a problem issue that can stimulate testees using higher order thinking skills (HOTS). c. Each text contains a theme that corresponds to the Indonesian indicator. d. Each text provides information that can add insight to science and the value of life to testees. e. Each text has a readable level.
<p style="text-align: center;">Question</p> <div style="border: 1px solid black; padding: 10px; margin: 10px auto; width: 80%;"> <div style="border: 1px solid black; padding: 5px; margin-bottom: 10px; text-align: center;">Stem</div> <div style="border: 1px solid black; height: 40px; width: 100%;"></div> </div>	<ul style="list-style-type: none"> a. The question stem is a question on the text that is read by testee. b. The stem contains questions that stimulate the testees using higher order thinking skills (HOTS) to answer questions. c. There are 5 answer choices. The number of alternative answers is sequential. d. Testee were given stimulation to analyze the questions more sharply with a more varied choice of answers. e. Each answer choice is made parallel. f. Each answer choice is made in a phrase or clause with correct spelling and accordance with the writing rules.

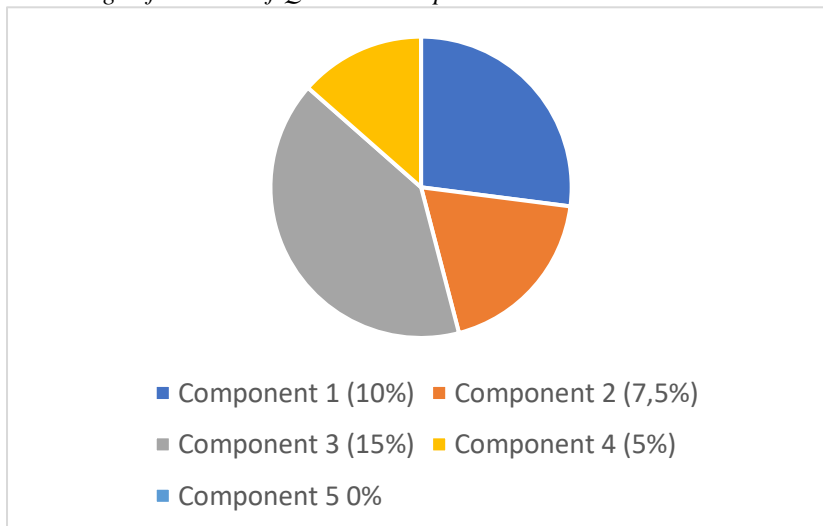
The proof of product accuracy could be seen from the results of data processing at the design and develop stages. In the design phase, the data obtained was the result of an expert judgment on the product. Products consisted of three sets, namely SET A, SET B, and SET C, which were analyzed through the five component assessment parameters. The description of the results of expert assessment of the product can be summarized in the following points.

- a. The context of Indonesian culture has been incorporated;
- b. HOTS degree should be considered;
- c. The relationship between text and questions should be accommodated;
- d. Questions of who and what need to be avoided;
- e. Technical writing questions should pay attention to the use of sentence writing rules;

- f. It is necessary to pay attention to the technical use of punctuation, lack of letters in words, and prepositions with capital letters; and
- g. It is necessary to simplify the sentence in the stem and in the choices that are too long.

The results of the qualitative assessment from the expert showed that the products developed in principle had fulfilled the requirements. A key area that needed to be improved pertained to the rules of writing and the rules of language. The aspect of material accuracy did not require much improvement. For more details, Figure 3 displays the percentage of the number of questions with improvement responses from the experts.

Figure 3
Percentage of Number of Questions Responded



Based on the expert opinion, a revision process was conducted. Afterwards, in the developing phase, the product of this study was empirically tested on the students to obtain data on the results of product implementation. Each set of measuring device products amounted to 25 questions, with a total of 75 questions. On the basis of data processing, all the questions were valid and reliable. The results of testing the validity can be seen in Table 4.

Table 4 shows the results of the calculation and processing of the 75 items in questions a, b and c showing that 75 items exhibit a significant validity index at $p < 0.05$. This means that all items fall into the valid category. In other words, the instruments can be used to measure reading literacy skill. In addition to the validity test, each set of questions was tested for reliability. The reliability test results can be seen in the following Table 5.

Based on the reliability coefficient category using Drummond and Jones (2010) classification, Table 5 shows the Cronbach Alpha value which represents the quality of the items in question A of 0.56 including the moderate category. In question B the quality of the items is 0.623 including the high category. In question C the quality of the items was 0.547 including the medium category. The validity and reliability test results demonstrate that the research product produced meets the requirements as a standardized reading literacy assessment product. Statistical and psychometric interpretations, such as the calculation of validity standards and reliability standards were used to accurately interpret assessment instruments (Denton et al., 2011; Webb, 2002). From the data obtained, this study generates several findings. Using HOTS-based reading comprehension parameters, this study furnishes evidence as shown in Figure 4.

Table 4
 Test Results for the Validity of HOTS-Based Reading Literacy Measurement Instruments in a Cultural Context

VALIDITY TEST RESULTS QUESTION A				VALIDITY TEST RESULTS QUESTION B				VALIDITY TEST RESULTS QUESTION C			
Numb.	r stat	r table	Criteria	Numb.	r stat	r table	Criteria	Numb.	r stat	r table	Criteria
1	0.202158	0.1678	valid	1	0.365261	0.1678	valid	1	0.3190787	0.1678	valid
2	0.230025	0.1678	valid	2	0.410746	0.1678	valid	2	0.4144884	0.1678	valid
3	0.190993	0.1678	valid	3	0.406843	0.1678	valid	3	0.3871607	0.1678	valid
4	0.405442	0.1678	valid	4	0.452984	0.1678	valid	4	0.5089796	0.1678	valid
5	0.173154	0.1678	valid	5	0.337924	0.1678	valid	5	0.3585744	0.1678	valid
6	0.307226	0.1678	valid	6	0.337961	0.1678	valid	6	0.1834442	0.1678	valid
7	0.228398	0.1678	valid	7	0.312371	0.1678	valid	7	0.5746154	0.1678	valid
8	0.256318	0.1678	valid	8	0.382613	0.1678	valid	8	0.4893802	0.1678	valid
9	0.195437	0.1678	valid	9	0.189227	0.1678	valid	9	0.435872	0.1678	valid
10	0.318121	0.1678	valid	10	0.211524	0.1678	valid	10	0.3153399	0.1678	valid
11	0.268946	0.1678	valid	11	0.432264	0.1678	valid	11	0.374397	0.1678	valid
12	0.44702	0.1678	valid	12	0.32615	0.1678	valid	12	0.1895331	0.1678	valid
13	0.230718	0.1678	valid	13	0.345643	0.1678	valid	13	0.5173624	0.1678	valid
14	0.412373	0.1678	valid	14	0.353733	0.1678	valid	14	0.3444105	0.1678	valid
15	0.224873	0.1678	valid	15	0.353233	0.1678	valid	15	0.2093623	0.1678	valid
16	0.336769	0.1678	valid	16	0.330326	0.1678	valid	16	0.2745626	0.1678	valid
17	0.282539	0.1678	valid	17	0.269324	0.1678	valid	17	0.3972563	0.1678	valid
18	0.261638	0.1678	valid	18	0.402294	0.1678	valid	18	0.3139718	0.1678	valid
19	0.25385	0.1678	valid	19	0.189627	0.1678	valid	19	0.1960238	0.1678	valid
20	0.456556	0.1678	valid	20	0.280488	0.1678	valid	20	0.2329172	0.1678	valid
21	0.332258	0.1678	valid	21	0.265374	0.1678	valid	21	0.323647	0.1678	valid
22	0.237097	0.1678	valid	22	0.309891	0.1678	valid	22	0.3839205	0.1678	valid
23	0.233676	0.1678	valid	23	0.214445	0.1678	valid	23	0.4701617	0.1678	valid
24	0.173275	0.1678	valid	24	0.230712	0.1678	valid	24	0.4488779	0.1678	valid
25	0.414696	0.1678	valid	25	0.177323	0.1678	valid	25	0.2111863	0.1678	valid

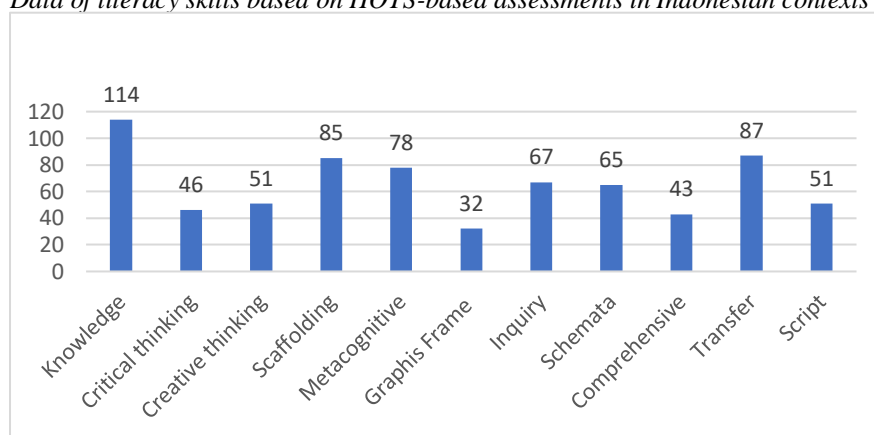
Table 5

Results of the Validity Test of HOTS-Based Reading Literacy Measuring Instruments in a Cultural Context

Reliability Test Results Question A			Reliability Test Results Question B			Reliability Test Results Question C		
	Range	Total		Range	Total		Range	Total
Reliability	0.392245	0.563471	Reliability	0.4528812	0.623425	Reliability	0.3743098	0.5447241
r table	0.1678	0.287378	r table	0.1678	0.287378	r table	0.1678	0.287378
Criteria	Reliable	Reliable	Criteria	Reliable	Reliable	Criteria	Reliable	Reliable

Figure 4

Data of literacy skills based on HOTS-based assessments in Indonesian contexts



This research reveals that based on the Reading Literacy Assessment Instrument developed in this study, the highest literacy reading skill is the ability of knowledge. That is, the ability most mastered by students is the ability to identify and remember factual data from a text. This is the ability with the lowest level. This Reading Literacy Assessment Instrument also suggests that junior high school students have not yet reached a high level of critical and creative reading. One of the efforts that can be done is to train students to think critically and creatively by solving reading questions based on HOTS. Among other abilities, both of these abilities appear to remain poor. This shows that junior high school students in Indonesia do not yet have higher order thinking skills in understanding and dealing with reading problems.

However, different from the results of other literacy instruments, the literacy tool under investigation indicates that the level of reading literacy skill of the middle school students under examination is not at a very low level. The reading literacy skills of junior high school students, in the domain of critical reading skills and creative reading skills, are close to achieving the expected 50% ability. Critical and creative cognitive abilities are at a high and complex level of cognitive hierarchy (Noble, 2004). The data above shows that the reading literacy level of middle school students in Indonesia is not apparently in an alarming situation. The Reading Literacy Assessment tool developed

has accurate readability measurements, the right context, and the content in line with the characters of the Indonesian nation. This assessment tool has not been tested extensively. Therefore, with a broader test it is expected that this instrument can show a more factual state of the level of reading literacy skill of junior high school students in Indonesia.

CONCLUSION

Reading literacy assessment is an important part of learning decision making. The reading literacy model proposed here was evidently valid and reliable, hence a potentially standardized reading tool to measure students' reading skills in the Indonesian context. With the production of these standardized reading literacy measures, the teacher can use them to provide a more thoughtful and meaningful assessment of reading literacy to students. Since the instrument developed was based on cognitive taxonomy with a complex hierarchy, it can stimulate students to enter into high-level critical and creative cognitive processes.

The recommendations proposed from the results of this study are as follows. Theoretically, there is a need to develop reading literacy instruments based on higher order thinking skills (HOTS) for effective and efficient students at the elementary and high school levels. Practically, the need exists for instruments that can be used

practically with easy procedures to measure reading literacy skills based on HOTS. In terms of policy, it is necessary to make a policy to more broadly test a HOTS-based reading literacy assessment in Junior High School in Indonesia.

ACKNOWLEDGEMENTS

This work was supported by Direktorat Riset dan Pengabdian kepada Masyarakat, Kementerian Riset dan Teknologi, Republik Indonesia.

REFERENCES

- Abidin, Y. (2013). *Pengembangan model penilaian autentik dalam pembelajaran membaca pemahaman di sekolah dasar* [Developing of an authentic assessment model in reading comprehension learning at primary school] [Unpublished doctoral dissertation]. Universitas Pendidikan Indonesia.
- Alonzo, J., Basaraba, D., Tindal, G., & Carriveau, R. S. (2009). They read, but how well do they understand?: An empirical look at the nuances of measuring reading comprehension. *Assessment for Effective Intervention*, 35(1), 34-44. <https://doi.org/10.1177%2F1534508408330082>
- Anderson, R. C. (1985). *Becoming a nation of readers: The report of the commission on reading*. West Virginia University Press.
- Barkan, S. E. (2011). *Sociology: understanding and changing the social world*. Flat World.
- Barton, D., Hamilton, M., & Ivanic, R. (Eds.). (2000). *Situated literacies reading and writing in context*. Routledge.
- Benjamin, A., & Hugelmeyer, M. (2013). *Big skills for the common core: Literacy strategies for the 6–12 classroom*. Routledge.
- Brookhart, S. M. (2010). *How to assess higher-order thinking skills in your classroom*. ASCD.
- Cole, M. (1998). *Cultural psychology: A once and future discipline*. Harvard University Press.
- Crandall, B. R., Lewis, E., Stevens, E. Y., Robertson, J. M., O'Toole, J. E., Cullen, K. A., Munger, K. A., Murray, M. S., Duffy, M. A., Dussling, T. M. & McQuitty, V. (2016). *Steps to success: Crossing the bridge between literacy research and practice*. Suny Open Textbooks. <https://textbooks.opensuny.org/steps-to-success/>
- Crowl, T. K., Kaminsky, S., & Podell, D. M. (1997). *Educational psychology: Windows on teaching*. Brown and Benchmark.
- Davies, B., & Bansel, P. (2007). Neoliberalism and education. *International Journal of Qualitative Studies in Education*, 20(3), 247-259. <https://doi.org/10.1080/09518390701281751>
- Denton, C. A., Ciancio, D. J., & Fletcher, J. M. (2011). Validity, reliability, and utility of the observation survey of early literacy achievement. *Reading Research Quarterly*, 41(1), 8-34. <https://doi.org/10.1598/RRQ.41.1.1>
- DeVries, R., & Kohlberg, L. (1987). *Programs of early education: The constructivist view*. Longman.
- Drummond, R.J. & Jones, K.D. (2010). *Assessment procedures for counselors and helping professionals* (7th ed.). Pearson.
- Edglossary (2014, April 30). Criterion-referenced test [online]. *The Glossary of Education Reform*. Retrieved from <http://edglossary.org/criterion-referenced-test/>
- Facione, P. A. (1998). *Critical thinking: What it is and why it counts*. California Academic Press.
- Florida Department of Education. (1996–1997). *The basics of school improvement and accountability in Florida*. Florida Department of Education.
- Fulcher, G. & Davidson, F. (2007). *Language testing and assessment: An advanced resource book*. Routledge.
- Frankel, K. K., Becker, B. L. C., Rowe, M. W., & Pearson, P. D. (2016). From "what is reading?" to what is literacy?. *Journal of Education*, 196(3), 7-17. <https://doi.org/10.1177%2F002205741619600303>
- Greenleaf, C. L., Litman, C., Hanson, T. L., Rosen, R., Boscardin, C. K., Herman, J., Schneider, S. A., Madden, S., & Jones, B. (2010). Integrating literacy and science in biology: Teaching and learning impacts of reading apprenticeship professional development. *American Educational Research Journal*, 48(3), 647-717. <https://doi.org/10.3102%2F0002831210384839>
- Gunes, F. (1997). *Okuma yazma ogretimi ve beyin teknolojisi*. Ocak Publications.
- Huot, J. (1995). Dimensions of learning. *Pedagogy & Curriculum Development*. <http://www.collegequarterly.org/CQ.html/HH.H.033.CQSpr95.Huot.html>
- International Reading Association & National Council of Teachers of English. (2010). *Standards for the assessment of reading and writing*. Newark. IRA & NCTE. <https://www.literacyworldwide.org/docs/default-source/resource-documents/standards-for-the-assessment-of-reading-and-writing.pdf>
- Koentjaraningrat. (1988). *Manusia dan kebudayaan di Indonesia* [Human and Culture in Indonesia]. Djambatan.

- Lapp, D., Moss, B., Grant, M., & Johnson, K. (2015). A close look at close reading: Teaching students to analyze complex texts, grades K-5. ASCD.
- Lau, K. C. (2009). A critical examination of PISA's assessment on scientific literacy. *International Journal of Science and Mathematics Education*, 7, 1061. <https://doi.org/10.1007/s10763-009-9154-2>
- Lewis, A. & Smith, D. (1993). Defining higher order thinking. *Theory into Practice*, 32(3), 131-137. <https://doi.org/10.1080/00405849309543588>
- Martin-Jones, M., & Jones, K. (Eds.). (2000). *Multilingual literacies*. John Benjamin Publishing Company.
- McQueen, J. & Mendelovits, J. (2003). PISA reading: Cultural equivalence in a cross-cultural study. *Language Testing Volume*, 20(2), 208-224. <https://doi.org/10.1191%2F0265532203lt2530a>
- McDavitt, D. S. (1994). Teaching for understanding: Attaining higher order learning and increased achievement through experiential instruction. ERIC Database No. ED374093 <https://files.eric.ed.gov/fulltext/ED374093.pdf>
- Morocco, C. C., Aguilar, C. M., Bershad, C., Kotula, A., & Hindin, A., (2008). *Supported literacy for adolescents: Transforming teaching and content learning for the Twenty-First Century*. Jossey-Bass A Wiley Imprint.
- Murray, M. S., Munger, K. A., & Clonan, S. M. (2011). Assessment as a strategy to increase oral reading fluency. *Intervention in School and Clinic*, 47(3), 144-151. <https://doi.org/10.1177%2F1053451211423812>
- Noble, T. (2004). Integrating the Revised Bloom's Taxonomy with multiple intelligences: A planning tool for curriculum differentiation. *Teachers College Record*, 106, 193-211. <http://dx.doi.org/10.1111/j.1467-9620.2004.00328.x>
- Ng, C., & Graham, S. (2017). Engaging readers in the twenty-first century: What we know and need to know more. In C. Ng & B. Bartlett (eds), *Improving reading and reading engagement in the 21st Century* (pp. 17-46). Springer. https://doi.org/10.1007/978-981-10-4331-4_2
- OECD (2016). *PISA 2015 results: Excellence and equity in education* (Volume I). OECD.
- Patrick, J. J. (1986). *Critical thinking in the social studies*. ERIC Clearinghouse for Social Studies/Social Science Education.
- Pearson, P. D., & Cervetti, G. N. (2013). The psychology and pedagogy of reading processes. In W. Reynolds & G. Miller (Eds.), *Educational psychology, Vol. 7: Handbook of psychology* (2nd ed., pp. 507-554). John Wiley & Sons.
- Picone-Zocchia, J. (2009). Changing the way you teach: Improving the way students learn. ASCD.
- Provost, M. C., Lambert, M. A., & Babkie, A. M. (2009). Informal reading inventories: Creating teacher-designed literature-based assessments. *Intervention in School and Clinic*, 45(4), 211-220. <https://doi.org/10.1177%2F1053451209353444>
- Purcell-Gates, V., Duke, N.K., & Stouffer, J. (2016). Teaching literacy: Reading. In D.H. Gitomer & C.A. Bell (Eds.), *The AERA handbook of research on teaching* (5th Ed., pp. 1217-1267). American Educational Research Association.
- Secretary's Commission on Achieving Necessary Skills (SCANS) (Ed.). (1991). *What work requires of schools: A SCANS report for America 2000*. Secretary's Commission on Achieving Necessary Skills, US Dept. of Labor.
- Sisson, D. & Sisson, B. (2014). *Close reading in elementary school: Bringing readers and texts together*. Routledge.
- Smagorinsky, P. (2009). The cultural practice of reading and the standardized assessment of reading instruction: When incommensurate worlds collide. *Educational Researcher*, 38(7), 522-527. <https://doi.org/10.3102/0013189X09347583>
- Snow, C. (2002). *Reading for understanding: Toward an R&D program in reading comprehension*. RAND Corporation
- Son, B., & VanSickle, R. L. (1993). *Problem-Solving instruction and students' acquisition, retention and structuring of economics knowledge*. ERIC Database No. ED366627. <https://files.eric.ed.gov/fulltext/ED366627.pdf>
- Sternberg, R. J., & Davidson, J. E. (Eds.). (1995). *The nature of insight*. MIT Press.
- Tantillo, S. (2012). *The literacy cookbook: A practical guide to effective reading, writing, speaking, and listening instruction*. John Wiley & Sons, Inc.
- TIMSS & PIRLS (2012). PIRLS 2011 international results in reading. TIMSS & PIRLS International Study Center.
- Trianto. (2011). *Mendesain model pembelajaran inovatif progresif* [Designing progressive innovative learning models]. Kencana Prenada Media Group.
- Trilling, B. & Fadel, C. (2009). 21st Century skills: Learning for life in our times. Jossey-Bass A Wiley Imprint.
- Vygotsky, L.S. (1978). Mind in society: The development of higher psychological

- process. Cambridge MA. Harvard University Press.
- Webb, N. (2002). *Assessment literacy in a standards-based urban education setting*. American Educational Research Association.
- Weeden, P., Winter, J., & Broadfoot, P. (2003). *Assessment: What's in it for schools?*. Routledge.
- Willis, J., Adie, L., & Klenowski, V. (2013). Conceptualising teachers' assessment literacies in an era of curriculum and assessment reform. *The Australian Educational Researcher, 40(2)*, 241-256. <https://doi.org/10.1007/s13384-013-0089-9>
- Wormeli, R. (2018). *Fair isn't always equal: Assessing & grading in the differentiated classroom*. Stenhouse Publishers.