IOSR Journal of Research & Method in Education (IOSR-JRME)
e-ISSN: 2320–1959.p- ISSN: 2320–1940 Volume 9, Issue 2 Ser. II. (Mar. - Apr .2019), PP 48-54
www.iosrjournals.org

Developing an Instrument of Science Literacy Assessment Based On the Local Context of Kalimantan for Junior High School Students

Choirul Amin[1], Abdullah[2], Arif Sholahuddin[2]
[1,2,3] Lambung Mangkurat University, Indonesia
Corresponding Author: Choirul Amin[1]

Abstract: *This study aims to develop an assessment instrument of Science Literacy (SL) for SMP students based on valid, practical, and effective paper also criteria that mostly use local context of Kalimantan. The methods used in this study was research & development (R & D) which adopted from Borg and Gall model. Subjects of the trial were IX grade students of five junior high schools in Banjarmasin which a total of 238 students. Preliminary field testing used six students from one school, main field testing used 50 students from one school and operational field testing used 182 students from three school. The results show that the assessment instrument has valid in terms of content and construction through expert review, and empirically valid with an average of validity index is 0.439. The instrument was also considered as reliable with reliability index dichotomous items is 0.611 and polytomous items is 0.734. The practicability of the instrument in terms of ease to use through three main stages, those are preparation, implementation, and analysis of assessment result. The effectivity in terms of the assessment objectives that has achieved is the results of the assessment obtained from main field testing and operational field testing which are relatively consistent showing the performance of the student's science literacy low level.*

Keywords: *Assessment instrument, science literacy*

Date of Submission: 24-03-2019 Date of acceptance: 08-04-2019

I. Introduction

Current curriculum of science teaching and learning for Indonesian junior high school uses an integrated concept that integrates knowledge of biology, physics, chemistry, earth and space. This integrated concept is shown in the core competencies and basic competencies in which learning practices are oriented toward applicative abilities, development of the ability to think, care, curiosity, and be responsible for the natural and social environment (Kemendikbud, 2013). Those characters of curriculum are in line with the concept of Science Literacy (SL).

Historically, SL is a term that specifically related to natural science which has impact to the social life. It is written in Hurd's (1958) article in Amerika which described clearly about the improvement of science and technology in chemistry, physics, biology, and earth-space science at that time. However nowadays, the definition of SL has been expanded. It does not only accommodate natural science field, but it is also broadly related to other science field. That is why the term broadly become scientific literacy. Until now, SL is still a hot issue in a country, even in America (Tomovic dkk, 2017).

PISA defines scientific literacy for assessing students' science performance as ability to engage with science-related issues, and with the ideas of science (OECD, 2013). In general, result of PISA's assessment shows that the level of scientific literacy of Indonesian students is still low. The data score from 2000 to 2012 are 393, 395, 393, 383, and 382 (OECD, 2003; OECD, 2004; OECD, 2007; OECDb, 2010; OECD, 2014). It is in line with some results of independent research which also show students low level of science literacy in Indonesia (Rachmatullah, Diana, & Rustaman, 2015; Aryani, Suwono, & Parno, 2016; Naturasari, Roshayanti & Nurwahyunani, 2016; Rusilowati, Kurniawati, Nugroho & Widiyatmoko, 2016; Sumaryatun, Rusilowati & Nugroho, 2016; Nofiana & Julianto, 2017; Siagian, Silitonga & Djulia, 2017; Andriani, Saparini & Akhsan, 2018; Nur'aini, Rahardjo & Susanti, 2018; Rahmadani, 2018). However, in the last PISA's assessment, it has experienced a significant improvement, reaching 403 (OECDa, 2016). This significant improvement of SL score proves that the change of curriculum has a positive influence. The new K13 curriculum has fulfills the criteria for improving student's SL (Wasis, 2015; Mustika & Rahmat, 2014; Anjarsari, 2014).

Responding to the 2012 PISA's report, Wasis (2015) said that 60% of junior high school students in Indonesia were only able to answer questions or problems accompanied by complementary supporting information, could identify information, but to use that information procedurally still needed clear direction and no convoluted, besides that they were also only able to take an action if only given a clear stimulus. While the data in 2015 explained that Indonesian students who have the best performance could use abstract scientific ideas or concepts to explain phenomena/events that are not familiar and more complex. While low-performance students were only able to use basic scientific knowledge to interpret data and explain valid scientific conclusions (OECDb, 2016).

However, PISA instrument for science assessment still has limitation. It is impossible to represent entire regional background and curriculum of each country. Therefore, science literacy assessment instrument needs to be developed in accordance with each respective local characteristic, but still refer to the assessment framework of PISA.

The success of K13 must be accompanied by teachers' skills in making assessment items that are in accordance with the characteristics of each local context of regional and country. The selection of local contexts which related to science content knowledge in the assessment will assess the competency that students have acquired in school. It also allows the students to understand the application of the knowledge that they encounter in daily life, to evaluate the application of the knowledge scientifically, and to interpret data and evidence they found. Therefore, the research and development of this assessment instrument mostly uses the local contexts of Kalimantan.

II. Method

The research and development (R&D) model refers to Borg and Gall which is defined as process or method for validating and developing educational products (Sugiyono, 2015). This model consists of ten steps: (1) research and information collecting, (2) planning, (3) develop preliminary form a product, (4) preliminary field testing, (5) main product revision, (6) main field testing, (7) operational product revision, (8) operational field testing, (9) final product revision, (10) dissemination and implementation. However, for this study is only limited to the final product revision because of our limited resources.

After preliminary product had developed, then it was tested by 5 experts to know the content and construct validity. Assessment criteria refers to four SL components and the legibility of each items as shown in Table 1. Assessment result was measured by using Content Validity Ratio (CVR) in three criteria, those are: (1) not necessary, (2) useful but not essential, (3) essential (Cohen & Swerdlik, 2009).

Table 1. Four Scientific Literacy Component & Legibility

Assessment Component for Expert Review	Description	
Competency	1.	Explain phenomena scientifically;
	2.	Evaluate and design scientific enquiry;
	3.	Interpret data and evidence scientifically.
Knowledge	1.	Content knowledge;
	2.	Procedural knowledge;
	3.	Epistemic knowledge.
Cognitive demand	1.	High (H);
	2.	Medium (M);
	3.	Low (L).
Context	1.	Individual;
	2.	Local/National;
	3.	Global.
Legibility	Easy to understand	

After expert validation, the instrument that had been validated can be used in preliminary field testing. Subjects used in preliminary field testing were six students of IX grade from one junior high school. Subjects used in main field testing were two classes consist of 50 students from one junior high school, and for operational field testing used each 2 classes from three schools consist of 182 students of IX grade.

Then, the obtained data was analyzed statistically and qualitatively based on 3 agreed parameters, those are validity, practicability, and effectivity. Items validity were measured using Pearson formula. Reliability was also analyzed using single-test-single-trial method. Item reliability for dichotomous questions was calculated using Spearman-Brown formula and polytomous questions using Alpha-Cronbach formula. In addition, the difficulty and discrimination index was also measured (Arikunto, 2012). Criteria for difficulty and discrimination index is shown in Table 2.

Table 2. Criteria for Difficulty and Discrimination Index

Difficulty Index (P)		Discrimination Index (D)	
Range	**Criteria**	**Range**	**Criteria**
0,00-0,30	Difficult	Negative	Very poor
0,31-0,70	Medium	0,00-0,20	Poor
0,71-1,00	Easy	0,21-0,40	Moderate

Difficulty Index (P)		Discrimination Index (D)	
Range	Criteria	Range	Criteria
		0,41-0,70	Good
		0,71-1,00	Very Good

Last parameter which must be determined was criteria that used to measure student's SL achievement. the criteria refer to 5 level as shown in Table 3 below.

Table 3. Criteria of Students Science Literacy Achievement

Range (%)	Criteria
0-20	Very low
21-40	Low
41-60	Moderate
61-80	High/good
81-100	Very high/good

III. Result and Discussion

First, result of expert validation was shown that all items has valid in term of content and construct validity which each item's score was 2 (essential). Those validity score were obtained after being revised based on expert instructions. Before expert validation was conducted, the number of items were 30 and decreased into 29 items after being validated. Those 29 items declared as valid for field testing.

Table 4. Validated Items' Questions

Science Literacy Competencies	Items
Explain phenomena scientifically	1, 2, 5, 7, 10, 11, 12, 18, 24, 25, 29
Evaluate and design scientific enquiry	8, 14, 16, 17, 20, 22, 23, 26
Interpret data and evidence scientifically	3, 4, 6, 9, 13, 15, 17, 19, 21, 27, 28

There is an item which is number 17 that used to measure two different competencies. It is because that item required students to interpret a set of data, then to determine a piece of that data for proposing a way in solving a problem which is also a part of competency to evaluate and design scientific enquiry.

Based on the cognitive demand, Table 5 below shows that there are seven questions belong to low level, fourteen questions belong to medium level, and ten questions belong to high level.

Table 5. Cognitive Demand of Questions

Cognitive Demand	Items
Low (L)	1, 2, 7, 11, 24, 27, 28
Medium (M)	3, 4, 5, 8, 9, 12, 14, 15, 16, 17, 18, 23, 25
High (H)	6, 10, 13, 19, 20, 21, 22, 26, 29

In the PISA framework (OECD, 2013) it is explained that low cognitive level is ability to carry out a one-step procedure, for example recalling of a fact, term, principle or concept or locating a single point of information from a graph or table. Medium cognitive level is described as ability to use and apply conceptual knowledge to describe or explain phenomena, select appropriate procedures involving two or more steps, organize/display data, interpret or use simple data sets or graph. While, high cognitive level is described as ability to analyze complex information or data, synthesize or evaluate evidence, justify, reason given various sources, develop a plan or sequence of steps to approach a problem.

Empirical data for validity will show in Table 6 below. It is shown that all items are valid based on obtained data from operational field testing because all index validity is above the criteria (0,145) compared to the moment product table and the average has reached 0,439. Meanwhile, this instrument is also declared to be reliable with reliability index reaching 0,611 for dichotomous items and 0,734 for polytomous items.

Table 6. Items Validity, Difficulty and Discrimination Index

Item Number	Validity Index	Description	Difficulty Index (P)		Discrimination Index (D)	
			P	Criteria	D	Criteria
1	0,495	Valid	0,505	Medium	0,497	Good
2	0,377	Valid	0,220	Difficult	0,211	Moderate
3	0,604	Valid	0,566	Medium	0,514	Good
4	0,412	Valid	0,560	Medium	0,370	Moderate
5	0,257	Valid	0,038	Difficult	0,063	Poor

Item Number	Validity Index	Description	Difficulty Index (P)		Discrimination Index (D)	
			P	Criteria	D	Criteria
6	0,390	Valid	0,214	Difficult	0,276	Moderate
7	0,292	Valid	0,830	Easy	0,191	Poor
8	0,420	Valid	0,049	Difficult	0,043	Poor
9	0,401	Valid	0,143	Difficult	0,140	Poor
10	0,365	Valid	0,077	Difficult	0,171	Poor
11	0,406	Valid	0,242	Difficult	0,248	Moderate
12	0,500	Valid	0,016	Difficult	0,037	Poor
13	0,148	Valid	0,126	Difficult	0,036	Poor
14	0,444	Valid	0,082	Difficult	0,173	Poor
15	0,402	Valid	0,495	Medium	0,297	Moderate
16	0,394	Valid	0,302	Medium	0,309	Moderate
17	0,377	Valid	0,396	Medium	0,297	Moderate
18	0,377	Valid	0,071	Difficult	0,092	Poor
19	0,372	Valid	0,033	Difficult	0,055	Poor
20	0,506	Valid	0,099	Difficult	0,220	Moderate
21	0,586	Valid	0,126	Difficult	0,258	Moderate
22	0,476	Valid	0,308	Medium	0,387	Moderate
23	0,540	Valid	0,670	Medium	0,422	Good
24	0,646	Valid	0,132	Difficult	0,248	Moderate
25	0,465	Valid	0,016	Difficult	0,037	Poor
26	0,540	Valid	0,160	Difficult	0,031	Poor
27	0,680	Valid	0,187	Difficult	0,327	Moderate
28	0,502	Valid	0,335	Medium	0,379	Moderate
29	0,370	Valid	0,165	Difficult	0,255	Moderate

Data of difficulty and discrimination index above shows that in operational field testing there is one item belong to easy, nine items belong to medium, and 19 items belong to difficult. Meanwhile, based on discrimination index, there are 12 items belong to poor criteria, and other items belong to moderate criteria.

Difficulty index define as the result of the number of students who answer correctly divided by the total number of students. While, discrimination index is the difference between the proportion of high achiever students group who answer correctly with low achiever group students who answer correctly (Arikunto, 2012).

Data in the Table 6 proves that the low of discrimination index item was caused by those items which were too easy or too difficult for students. So, for item number seven, many students could answer that question even for students who had low achievement. However, other items are considered into difficult questions. It is because those questions could not be answered by most of students, even for students who had high achievement. But this phenomenon does not surprise us because based on some researches' results found that the scientific literacy level of students in Indonesia is still low. It means that the unfamiliarity to the kind of scientific literacy questions become main issue that causes a low index of discrimination. So, in this case, it is not wise to consider the quality of SL questions to be accepted or rejected because of the discrimination index only. However, other considerations are needed such as empirical validity, content and construct validity, reliability, and student response (misunderstanding or not).

Data for student's achievement in science literacy is shown in percentage based on three competencies which are explain phenomena scientifically, evaluate and design scientific enquiry, and interpret data and evidence scientifically. In order to know the instrument consistency, data was compared between main field testing and operational field testing.

According to the criteria in the Table 3, the Fig. 1 shows that students' performance in scientific literacy was relatively consistent in low level. Percentage result of data from main field testing higher than operational field testing because of samples used in operational field testing higher than main field testing and more heterogeneous in term of cognitive performance. This greater heterogeneity because of the samples were taken from three different schools, whereas in the main field testing only taken from one school.

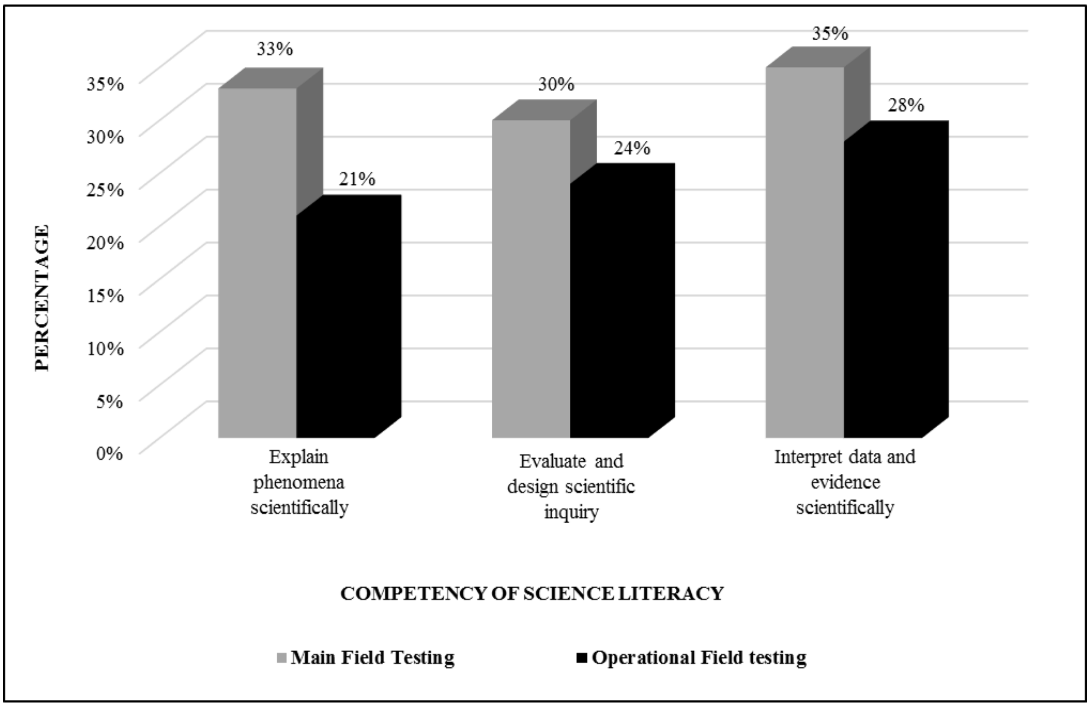

Figure 1. Students' Performance in Science Literacy

Data of cognitive demand is also displayed in Fig. 2 below and shown in percentage by comparing data obtained from main field testing and operational field testing. This Figure proves the consistency of obtained data that in general, the higher cognitive demand of item, the more difficult it is. Whether in main field testing or operational field testing, more than 50% students could not answer the question correctly. It is also mean that more students still did not able yet to use higher cognitive demand. Seemingly, teachers in particular need to familiarize their students by giving them questions that have high cognitive demands in their respective schools.

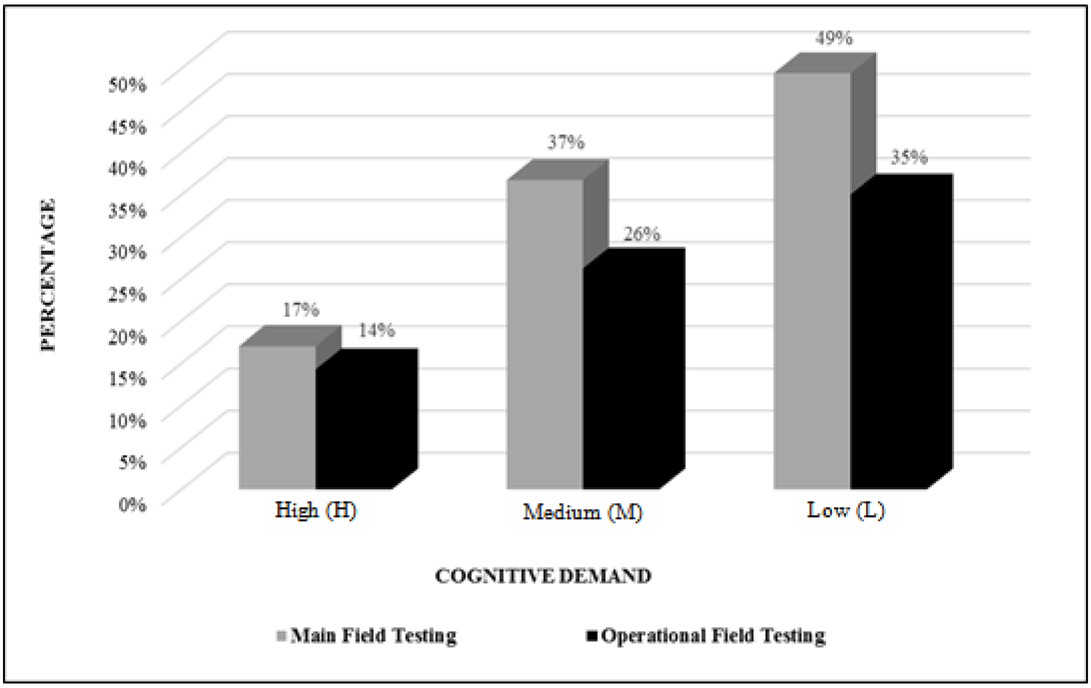

Figure 2. Students' Performance Based on Cognitive Demand

Scientific literacy questions have special character which constructed by scientific knowledge in real contexts based on facts and evidences which require students to explain phenomena scientifically, to evaluate and to design scientific enquiry, to interpret data and evidence scientifically. To explain phenomena, the students need ability to recall and apply scientific knowledge; to identify, use, and generate a model or representation; to offer hypotheses; and to make a prediction. To evaluate and design scientific enquiry, the students need ability to identify a problem that explored; to describe and evaluate a range of way used by scientists to ensure reliability and objectivity and generalizability explanation; to evaluate and to propose a way to explore a problem scientifically; also to distinguish problem that possible to investigate scientifically or not. To interpret data and evidence scientifically, the students need ability to identify assumption, evidence, and reasoning in scientific article; to analyze and interpret appropriate data; to transform data from one representation to others; to draw appropriate conclusion; to distinguish and evaluate scientific argument.

Based on the result of the students' respond, they looked familiar to the terms used in local context, but most of them unfamiliar to other knowledge, like how that local context works. For example, they know what peat land is, but they do not know that peat land has potential to save water and energy reserves, even how the damage of peat land is related to global warming, etc.

If we look at Table 5, item number seven is used to measure the low cognitive level and based on students' respond, the topic was familiar for them. It is different from items number one and two which are used to measure similar cognitive demand. So, the level of cognitive demand is not always in line with difficulty index. Familiarity of question's topic is another factor which influences difficulty index. It also influences the discrimination index.

According to the result of data analysis, there was at least seven reasons, "why did some items get bad score in discrimination index?".

1) The question had low level cognitive demand and students were familiar to the question, so most of students could answer the question;
2) The question had low level cognitive demand and students were unfamiliar to the question, so most of students could not answer the question;
3) The question had medium level cognitive demand but students did not carefully answer the question, so they did not answer the question perfectly;
4) The question had medium level cognitive demand, but students were unfamiliar to the content in the question;
5) The question had high level cognitive demand, but students did not carefully read the article or data in the question, even though the answer was written in it;
6) The question had high level cognitive demand, but students were unfamiliar in carrying out simple mathematic calculation;
7) The question had high level cognitive demand, but students were unfamiliar to the content in the question.

Term of familiarity in those seven reasons above is representing the level of students' knowledge (content, procedural, and epistemic knowledge).

Of course those seven reasons were also logically supported by PISA's report and independent researches that had conducted to know or find out the students' performance in scientific literacy. Most of their findings show that Indonesian students' performance is in low level.

IV. Conclusions

A range of procedure has been done to produce paper-based LS assessment instrument. All succeed criteria has been also reached based on validity, effectiveness, and practicability.

All items were declared as valid instrument based on content and construct validity in expert review or validation which the score of each items were 2 (essential). Each questions' item was also empirically considered as valid instrument based on data obtained from operational field testing as shown in Table 6 with the average score was 0,439 and it was also reliable with reliability index reaching 0,611 for dichotomous items and 0,734 for polytomous items.

Based on the data of cognitive demand and competency that compared between main field testing and operational field testing as shown in Figure 1 and 2, it proves that they were relatively consistent. In general, the higher cognitive demand of item, the more difficult it is, and the students' competency is relatively in low level. This is indicator which proves that the instrument of SL assessment is effective for assessing SL competency of junior high school students.

This instrument is also considered as practical assessment instrument which proved by data obtained from the application experience in assessing SL through 3 main procedures which were preparation stage, implementation stage, and score analysis.

References

[1]. Andriani, N., Saparini., Akhsan, H. (2018). Kemampuan Literasi Sains Fisika Siswa SMP Kelas VII Di Sumatera Selatan Menggunakan Kerangka PISA. *Berkala Ilmiah Pendidikan Fisika*. 6(3). 278-291.

[2]. Anjarsari, P. (2014). Literasi Sains Dalam Kurikulum dan Pembelajaran IPA SMP. *Prosiding Semnas Pensa VI* (602-607). Surabaya, Indonesia.

[3]. Arikunto. (2012). *Dasar-dasar Evaluasi Pendidikan Edisi 2*. Jakarta, Indonesia: PT Bumi Aksara.

[4]. Aryani, A, K., Suwono, H., Parno. (2016). Profil Kemampuan Literasi Sains Siswa SMPN 3 Batu. *Prosiding Semnas Pendidikan IPA Pascasarjana UM*. 1. 847-855.

[5]. Cohen, R, J., & Swerdlik, M. (2009). *Psychological Testing and Assessment: An Introduction to Test and Measurement 7th Edition*. New York: McGraw-Hill.

[6]. Hurd, P. (1958). Science Literacy: Its Meaning for American Schools. *Educational Leadership* (13). America.

[7]. Kemendikbud. (2013). *Panduan Mata Pelajaran IPA SMP*. Jakarta, Indonesia: Kemendikbud.

[8]. Mustika, E., & Rahmat, P. (2014). Implementasi Kurikulum 2013 Berbasis Pendekatan Saintifik Dalam Mengembangkan Literasi ilmiah di Sekolah Dasar Kota Bekasi. *Prosiding Semnas Pensa VI* (106-111). Surabaya, Indonesia.

[9]. Naturasari, H., Roshayanti, F., Nurwahyunani, A. (2016). Profil Kualitas Literasi Sains Siswa SMP Se-Kabupaten Pati. *Bioma*. 5(2). 1-13.

[10]. Noviana, M., & Julianto, T. (2017). Profil Kemampuan Literasi Sains Siwa SMP di Kota Purwokerto Ditinjau dari Aspek Konten, Proses, dan Konteks Sains. *Jurnal Sains Sosial dan Humaniora*. 1(2), 77-84.

[11]. Nur'aini, D., Rahardjo, S, E., Susanti, V, H, E. (2018). Students Profile About Science Literacy in Surakarta. *Journal of Physics Conference Series 1022*. 1-6.

[12]. OECD. (2003). *Literacy Skills for The World of Tomorrow Further Result from PISA 2000.* Paris, Prancis: OECD.

[13]. OECD. (2004). *Learning for Tomorrow's World First Result from PISA 2003*. Paris, Prancis: OECD.

[14]. OECD. (2007). *PISA 2006: Science Competencies for Tomorrow's World Volume 1*. Paris, Prancis: OECD.

[15]. OECD. (2010). PISA 2009 Result: What Students Know and Can Do, Student Performance in Reading, Mathematics and Science Volume 1. Paris, Prancis: OECD.

[16]. OECD. (2013). *PISA 2015 Draft Science Framework*. Paris, Prancis: OECD.

[17]. OECD. (2014). *PISA 2012 Result in Focus What 15-Years-Olds Know and What They Can Do With What They Know*. Paris, Prancis: OECD.

[18]. OECDa. (2016). *PISA 2015 Result in Focus*. Paris, Prancis: OECD.

[19]. OECDb. (2016). PISA 2015: Country Overview. Taken from: https://www.compareyourcountry.org/pisa/country/IDN?lg=en

[20]. Rahmadani, Y., Fitakurahmah, N., Fungki, N., Prihatin, R., Majid, Q., Prayitno, B, A. (2018). Profil Keterampilan Literasi Sains Siswa Sekolah Menengah Atas (SMA) di Karanganyar. *Jurnal Pendidikan Biologi*. 7(3). 183-190.

[21]. Rachmatullah, A., Diana, S., Rustaman, N, Y. (2015) Profile of Middle School Students on Scientific Literacy Achievements by Using Scientific Literacy Assessments (SLA). *AIP Conference Proceedings*. 1-5.

[22]. Rusilowati, A., Kurniawati, L., Nugroho, S, E., Widiyatmoko, A. (2016). Developing an Instrument of Scientific Literacy Assessment on The Cycle Theme. *International Journal of Environmental & Science Education*. 11(12). 5718-5727.

[23]. Siagian, P., Silitonga, M., Djulia, E. (2017). Scientific Literacy Skills of Seventh Grade Junior High School (SMP Negeri) Students in North Labuhanbatu Regency. *International Journal of Humanities Social Sciences and Education (IJHSSE)*. 4(11). 176-182.

[24]. Sugiyono. (2015). *Metode Penelitian dan Pengembangan*. Bandung, Indonesia: Alfabeta.

[25]. Sumaryatun., Rusilowati, A., Nugroho, S, E. (2016). Pengembangan Instrumen Penilaian Autentik Kurikulum 2013 Berbasis Literasi Sains Pada Materi Bioteknologi. *Journal of Primary Education*. 5(1). 66-73.

[26]. Tomovic, C., McKinney, S., Berube, C. (2017). Scientific Literacy Matters: Using Literature to Meet Next Generation Science Standards and 21st Century Skills. *K-12 STEM Education*. 3(2). 179-191.

[27]. Wasis. (2015). Hasil Pembelajaran Sains di Indonesia: Problem dan Upaya Mengatasinya. *Prosiding Seminar Nasional Tahun 2015* (vi-xvi). Surabaya, Indoensia.

Choirul Amin. "Developing an Instrument of Science Literacy Assessment Based On the Local Context of Kalimantan for Junior High School Students". IOSR Journal of Research & Method in Education (IOSR-JRME) , vol. 9, no. 2, 2019, pp. 48-54.