

# Assessment of Literacy and Numeracy of Indonesian Students Using the Integrated Science-based Science Literacy Test

Purwo Susongko<sup>1</sup>  
{purwosusongko@upstegal.ac.id<sup>1</sup>}

Universitas Pancasakti Tegal<sup>1</sup>

**Abstract.** People's high literacy and numeracy levels are significantly influenced by the development of a nation. This study aims at developing the assessment instruments on Senior High School science program students' literacy and numeracy ability based on the integrated sciences. The literacy and numeracy indicators used were some indicators of the scientific and mathematics literacy achievements developed by PISA 2015. The literacy and numeracy test instruments were in the form of testlet consisting of 12 popular science readings containing 20 items measuring the literacy and 16 items measuring the numeracy. The instruments were tested to 310 XII grade students of SMAN 2 and SMAN 3 Tegal City. The test validity was examined using the Rasch model approach. The Senior High School science program students' literacy and numeracy test instruments have met the validity aspects both related to the material and psychometric aspects. The construct validity aspects with the Rasch model approach showed that the Senior High School science program students' literacy and numeracy test instruments were valid seen from the difficulty level and item appropriateness with the model. The literacy achievement of the XII Grade students of SMAN 2 and SMAN 3 Tegal City was 57.45 % while their numeracy achievement was 43.18 %.

**Keywords:** Assessment, literacy, numeracy, integrated mathematics and natural sciences

## 1 Introduction

People's high literacy and numeracy levels are significantly influenced by the development of a nation. People's scientific literacy positively influences the economic, democracy, cultural, individual development quality [1]-[3]. The future of a nation depends on its people with high scientific literacy levels to participate in various development fields as the society members possessing the information as well as the skillful scientific workforces who are ready to face the problematic challenges either in the local, national and global levels [4]. Therefore learning in schools must have an impact on improving students' literacy and numeracy skills. In order to find out the increase in these two aspects, it is necessary to develop an assessment model.

To achieve those literacy and numeracy achievements, the Republic of Indonesia Minister of Education and Culture, Nadiem Makarim starting in 2021 will change the National Examination with the minimum competency assessment and character survey. Nadiem has explained that what is meant by the minimum competency assessment refers to two elements: literacy and numeracy. The literacy meant here is not only related to the ability to but also the ability to analyze a text, ability to understand a concept behind the text. Furthermore, Nadiem

has also explained that numeracy is the ability to analyze numbers. Thus, the minimum competency assessment that will be performed is not based on the learning subjects but more on the mastery of contents or materials [5]. From those explanations, it can be concluded that the minimum competency assessment is not based on the basic competencies of each material given to the students, but based on the literacy and numeracy ability which are highly required in the working world [6],[7].

The scientific literacy achievement indicators developed by PISA (Programme for International Science Student Assessment) cover the competencies in: (1) scientifically explaining the phenomena, (2) scientifically interpreting the data and evidence, and (3) evaluating and designing the scientific investigations. Meanwhile, the mathematics literacy is defined as an individual ability to formulate, utilize, and interpret mathematics in various contexts including mathematical reasoning and mathematical concept utilization, procedure, fact, instrument to describe, explain, and predict a phenomenon [8]. Numeracy is the ability to access, utilize, interpret, communicate the mathematical information and ideas, to involve in the mathematical implementation from various life situations [9], [10]. Thus, there will be no differences between the definition of mathematics literacy and numeracy.

Due to the global competency demands, and the policy made by the Republic of Indonesia Minister of Education and Culture, it is necessary to arrange test instruments which have the ability to measure the literacy and numeracy aspects at all school levels. This study aims at developing the assessment instruments on the Senior High School science program Students' literacy and numeracy levels based on the integrated sciences covering (1) test construction, (2) test validity examination and (3) test result analysis. This study was a piloting project conducted at the Senior High Schools, especially for the Mathematics and Natural Sciences Program in Tegal City, Central Java, Indonesia. The use of Integrated Sciences Theme was selected based on the research result shown by the integrated Mathematics and Natural Sciences learning had a higher influence on the students' scientific literacy improvement [11], [12].

## 2 Method

The literacy and numeracy indicators used are some scientific and mathematics indicators developed by PISA 2015 as presented in Table 1.

**Table 1.** Scientific literacy Achievements and Indicators based on PISA 2015

<b>Dimension</b>	<b>Indicator used</b>
Literacy	Recalling and implementing the appropriate scientific knowledge. Identifying, utilizing, and ability to create an explanation model. Transforming data from one representation to the other representation. Properly analyzing and interpreting data as well as drawing conclusion. Identifying assumptions, evidence, and reasons in the texts related to science.
Numeracy	Implementing mathematics to create a generalization. Ability to access, utilize, interpret, and communicate the mathematical information in various contexts.

The test development stages for the Senior High School science program students' literacy and numeracy assessments were started by developing the indicators and themes of the integrated mathematics and natural sciences as well as determining the test form. The themes

were taken from various sources, such as Science News, [www.sciencenewsforstudents.org](http://www.sciencenewsforstudents.org), [www.readwork.org](http://www.readwork.org), and other sources. The selected test form was the testlet in which each testlet consisting of 3 items as the indicators of literacy and numeracy achievements. The test instruments consisted of 12 testlets tested to 310 science program students of SMAN 2 and SMAN 3 Tegal City in February 2020.

12 themes or materials were used in the literacy and numeracy test. Those 12 themes were popular science readings. Each theme has three items to measure the literacy and numeracy level that there were 20 items of literacy test and 16 items of numeracy test. The test instruments have been validated by two study field experts and two psychometric field experts as well as 10 teachers and practitioners.

In the modern test theory measurement, Rasch model is considered as the most objective measurement model. The utilization of Rasch model in education measurements has the excellence in high specific objectivity and item parameter estimation stability. Rasch model connects the opportunities to correctly answer each item ( $P(\theta)$ ) as the function of ability ( $\theta$ ) with the constant item difficulty level ( $b$ ) through the relationship presented in equation 1 [13].

$$P_i(\theta) = \frac{e^{(\theta - b_i)}}{1 + e^{(\theta - b_i)}} \quad (1)$$

Validity using the Rasch model refers to the Messick validity where construct validity is considered as a single concept consisting of some aspects [14]. Rasch analysis explained more comprehensive construct validity than the classical test theory. At least, there are six construct validity aspects: content, substantive, structural, external, generalizability, and consequential aspect [15]. The content validity aspect, related to the empirical evidence, mentioned that the test items which are appropriate or based on model are used as well as the items' difficulty levels are in accordance with the test participants' difficulty levels [16].

### 3 Results and Discussion

The validators' interview results showed that the literacy and numeracy test constructions were valid seen from the material, psychometric, and appropriateness aspects with the test subjects. The example of literacy and numeracy test items in the third theme is shown as follows. The heating oceans resulted from the climate changes produce less fish

It is harder to catch the fish due to the climate changes which continuously heat the world oceans. The oceans' increasing temperature for more than 80 years has continuously reduce the catching of 124 fish and shellfish species which can be harvested without causing long-term damages to the population up to 4.1 percent as reported by a recent study. Excessive catching has worsened the decrease, said the researchers. In some parts of the world, such as in Japanese Ocean where catching was excessively made, the catching decrease reached 35 percent. This study, on 1 March Science, the researchers investigated the changes starting from 1930 to 2010 on 235 fish and shellfish populations spread in 38 ocean areas. Averagely, the temperature of the earth ocean has increased approximately half degree Celsius at that time although the temperature changes varied from one location to the others.

Approximately 8 percent of fish and shellfish population investigated experienced loss caused by the ocean heating, while approximately 4 percent of populations increased at that time. Certain species, such as black sea bass along the east coast of the US Ocean, have grown very well in warmer waters. However, with the continuous heating, the benefits tended to evaporate and even those fish have reached their heat limits, said Christopher Free, a

quantitative ecological expert from the University of California, Santa Barbara, who led the project when he was at the University of Rutgers, New Brunswick, NJ.

1. It is explained in the passage that climate changes caused the heating water temperature of the ocean water surface. The followings are the chemical compounds available on the air making the earth temperature increase, except  
 A. CO<sub>2</sub>      B. CH<sub>4</sub>      C. O<sub>2</sub>      D. H<sub>2</sub>O      E. SO<sub>2</sub>      (2)
2. According to the passage, the followings are the appropriate explanations related to the relationship between the increasing ocean water temperature and the decreasing fish catching:
  - A. at high water temperature, the oxygen concentration will relatively decrease that many fish find them difficult to live and select the cooler water temperature
  - B. all fish cannot live at the medium or high water temperature
  - C. all fish prefer living at the extremely low water temperature
  - D. the increasing water surface temperature at the ocean makes the fish moves more actively that they are harder to catch
  - E. many fish died due to the increasing ocean water temperature
3. The earth ocean surface temperature has increased approximately half degree Celcius from 1930 to 2010. Based on the data, the temperature in 2130 will presumably increase up to  
 A. 0.5°      B. 1.0°      C. 1.5°      D. 2.0°      E. 2.5°      (3)

Those test item examples shows that the test item 3a and 3b measures the literacy ability, while test item 3c measures the numeracy ability. If the test participants truly understand the passage and have well mastered the scientific knowledge, the students will easily answer the test item 3a and 3b. Meanwhile, the test item 3c is related to the ability to predict based on the data explained in the passage. Students with good mathematical reasoning will easily answer the test item 3c.

By using the Rasch model, all test items which measure both literacy and numeracy ability have the difficulty level in accordance with the test participants' ability and have the appropriateness with the model used, as shown in Table 3 and Table 4. The ideal difficulty level is between -2 to 2. At those ranges with a normal distribution approach, 95.46 % of students' ability can be measured through the test.

**Table 2.** Test item difficulty level and appropriateness in the literacy test

Item	Difficulty Level	P-Value	P Outfit-t	Infit-t	Appropriateness
1a	0.214	0.498	-0.092	0.268	Appropriate
1b	- 2.135	0.999	-1.120	-0.703	Appropriate
2a	-1.631	0.999	-1.559	-0.683	Appropriate
2b	-0.502	0.984	-2.307	-1.767	Appropriate
3a	0.201	0.291	0.841	1.023	Appropriate
3b	0.283	0.727	-1.165	-1.553	Appropriate
4a	1.286	0.004	2.853	2.650	Appropriate
4b	2.385	0.023	1.044	-0.236	Appropriate
5a	0.951	0.826	-1.402	-1.203	Appropriate
5b	-0.407	0.522	-0.114	0.239	Appropriate
6a	-1.859	0.999	-1.339	-0.718	Appropriate
6c	0.447	0.443	0.151	0.131	Appropriate
7a	0.584	0.165	1.599	1.519	Appropriate

7c	-0.391	0.966	-2.142	-1.988	Appropriate
8a	-0.791	1.000	-2.895	-2.020	Appropriate
9a	-1.216	0.993	-1.569	-1.214	Appropriate
9c	2.328	0.002	1.575	0.417	Appropriate
10a	0.488	0.274	0.968	1.112	Appropriate
11a	0.807	0.378	0.399	0.475	Appropriate
12a	-1.042	0.955	0.865	0.922	Appropriate

**Table 3.** Test item difficulty level and appropriateness in the numeracy test

Item	Difficulty Level	P-Value	P Outfit-t	Infit-t	Appropriateness
1c	-1.561	0.652	-0.355	-0.495	Appropriate
2c	0.339	0.986	-2.672	-2.551	Appropriate
3c	-0.142	0.345	0.554	0.882	Appropriate
4c	-0.593	0.652	-0.728	-0.508	Appropriate
5c	-0.524	0.733	-1.134	-0.985	Appropriate
6b	-2.081	0.780	-0.454	-0.092	Appropriate
7b	0.309	0.984	-2.671	-2.743	Appropriate
8b	0.874	0.612	-0.251	-0.079	Appropriate
8c	-0.073	0.013	3.390	3.455	Appropriate
9b	2.025	0.000	2.294	0.462	Appropriate
10b	0.750	0.783	-0.716	-1.057	Appropriate
10c	0.683	0.959	-1.644	-1.295	Appropriate
11b	0.236	0.165	1.188	2.003	Appropriate
11c	-0.210	0.539	-0.256	-0.202	Appropriate
12b	0.354	0.774	-0.941	-1.455	Appropriate
12c	-0.387	0.765	-1.325	-1.402	Appropriate

Minimally, there were two criteria in viewing the test item appropriateness with the model: by seeing the opportunity price of data appropriateness test with the model and both outfit-t and infit-t values. By taking the significance level of 0.01, the p-value for each item having the value higher than 0.01 was considered appropriate with the model. Meanwhile, both outfit-t and infit-t values were considered still appropriate with the value model of between -2 and 2, for big samples could be tolerated from -3 to 3. In the literacy items, there were two items having the p-value of less than 0.01, yet from both outfit-t and infit-t values were still appropriate with the model. Meanwhile, in numeracy items, all items had the p-value of more than 0.01 that it can be concluded that those were appropriate with the Rasch model used.

The literacy scores obtained by the students were at the maximum level of 18, while the minimum score was 5 with the students' average achievement of 11.5 or 57.45 % from the maximum score. The literacy score interval was 13, while the literacy score standard deviation was 2.6. The students' numeracy scores were at the maximum score of 12 and the minimum score of 0 with the students' average achievement of 6.9 or 43.18% from the maximum score. The numeracy interval score was 12, while the numeracy score standard deviation was 2.3. The results of the analysis also show that literacy and numeracy skills are correlated by 0.46 at a significance level of 5%. The results of this study are in line with previous studies that show literacy and numeracy abilities have a significant correlation [17],[18].

## 4 Conclusion

The Senior High School science program students' literacy and numeracy test instruments met the validity aspects both from the material and psychometric aspects. The construct validity aspect with the Rasch model approach showed that the Senior High School science program students' literacy and numeracy test instruments were valid seen from the level of difficulty and appropriateness of items with the model. The literacy achievement of the XII Grade Students of SMAN 2 and SMAN 3 Tegal City was 57.45 % and their numeracy achievement was 43.18 %. From the results of this study, this instrument can be used to assess literacy and numeracy skills at the senior high school level of the Sciences program.

## References

- [1] Hanushek EA, & Woessmann, L 2016 *Science* 351 6271 344-345
- [2] Rudolph JL, & Horibe S 2016 *Journal of Research in Science Teaching* 53 6 805-820.
- [3] Bereiter C 2002 *Cognitive Studies* 9 3 321-327.
- [4] National Assessment Governing Board 2008 *Science Framework for the 2009 National Assessment of Educational Progress* Washington DC: US Government
- [5] Nugraheny DE 2019 *Mendikbud: Penilaian Kompetensi untuk Pengganti UN Dilakukan di Kelas 4, 8, dan 11*. Artikel dalam *Harian Kompas*, 11-12-2019
- [6] Shomos A. & Forbes M 2014 *Literacy and numeracy skills and labour market outcomes in Australia*. Australian Government Productivity Commission.
- [7] Arriagada P & Hango D 2016 *Literacy and Numeracy among Off-Reserve First Nations People and Métis: Do Higher Skill Levels Improve Labour Market Outcomes? Insights on Canadian Society*. Statistics Canada.
- [8] OECD 2016 *PISA 2015 Assessment and Analytical Framework: Science, Reading, Mathematic and Financial Literacy* Paris : OECD Publishing
- [9] European Commission/EACEA/Eurydice 2015 *Adult Education and Training in Europe: Widening Access to Learning Opportunities* Luxembourg: Eurydice Report Publications Office of the European Union
- [10] Gal I. & Tout D 2014 *Comparison of PIAAC and PISA frameworks for numeracy and mathematical literacy*
- [11] Tamassia L, & Frans R 2014 *Journal of the European Teacher Education Network* 9 131-141
- [12] Maria A 2008 *Defining Integrated Science Education and Putting It to Test*. The Swedish National Graduate. Tidak dipublikasikan
- [13] Wu M & Adams R 2007 *Applying the Rasch model to psycho-social measurement: A practical approach*. Melbourne: Educational Measurement Solutions
- [14] Ravand H & Firoozi T 2016 *International Journal of Language Testing*, 6 11-18.
- [15] Wang C C, Ho HC., Cheng CL & Cheng YY 2014 *Creativity Research Journal* 26 1 62-71
- [16] Runnels J 2012 *International Journal of Language Studies* 6 4 141-155.
- [17] Davidse NJ, De Jong MT & Bus AG 2014 *Reading and Writing* 27 4 631-648.
- [18] Neumann MM, Hood M, Ford RM, & Neumann DL 2013 *European Early Childhood Education Research Journal* 21 4 489-501.