A case study of scientific literacy in natural science subject using rasch analysis model (RAM)

N Mubarokah¹, A Permanasari², and E Eliyawati*¹

¹International Program on Science Education Study Program, Universitas Pendidikan Indonesia, Dr Setia Budhi Street No. 229, Bandung 40154, Indonesia
²Department of Chemistry Education, Universitas Pendidikan Indonesia, Dr Setia Budhi Street No. 229, Bandung 40154, Indonesia

*eliyawati@upi.edu

Abstract. The study aims to analyze junior high school students’ scientific literacy in Natural Science subjects. The study used a case study method. There were 70 students of IX grade in a private junior high school in West Bandung Regency. Respondents consisted of 40 boys and 30 girls with an average age of 15 years old. The instrument used was a written test consisting of 30 simple multiple-choice questions for three aspects of scientific literacy, which were scientific knowledge, scientific competencies, and attitudes toward science. The data analysis used Rasch Analysis Model (RAM) and assisted by Winsteps 4.4.5. The results being discussed are the person and item reliability, variable maps, and differential item functioning (DIF). The results show that the average scientific literacy ability of students falls into the low category, and some questions have gender biases. Thus, scientific literacy needs to be promoted and involved effectively in teaching-learning activities to get students used to the scientific literacy environment in their learning programs during school years.

1. Introduction
Modern societies encounter societal issues and discuss solutions at the global stage [1]. People may have to deal with many other cases of decision-making that will have a significant impact on their safety and well-being, such as the efficient use of goods and technology and regular energy consumption [2], [3]. These examples are merely a reflection of the choices people need to make on scientific issues, and thus indicate the urgency of promoting scientific literacy for each individual. In the beginning of the late twenty-first century, Organization for Economic Cooperation and Development (OECD), the European Union (EU), and the United States (US) had identified and analyzed the competencies that future citizens will need to face global competitions [4]. Global competition refers to the intense competition of skills in the local, national, and international contexts.

Scientific literacy is an essential key to face various issues in the 21st century. Citizens of the 21st century need to learn the basics of science and technology. Science and technology advancement is a great potential to enhance the well-being of individuals, increase economic growth, and promote social equality, along with many other benefits [5]. Science literacy helps people to learn to become responsible citizens and to be fully aware of global problems.

Students’ development of scientific literacy involves the knowledge of science, the procedures of science, the scientific attitude, and the proficiency of science. Students, therefore, may not only understand the concepts of science but may also build the habit of implementing scientific ways to solve
problems by making informed choices based on scientific considerations [6]. According to a science assessment framework published by OECD, which was used for Programme for International Student Assessment (PISA) in 2015, there are three aspects of scientific literacy, which are scientific knowledge, scientific competencies, and attitudes toward science. Scientific knowledge consists of Content Knowledge, Procedural Knowledge, and Epistemic Knowledge. Scientific competencies include Explain Phenomena Scientifically, Evaluate and Design Scientific Inquiries, and Interpret Data and Evidence Scientifically. Attitudes toward science are Interest in Science and Technology, Valuing Scientific Inquiries, and Environmental Awareness [7].

Indonesia has been participated in the assessment from 2000 until the last assessment, which was held in 2018. Most of the time, Indonesian students show decreasing score results in the science domain. The scores are also below the OECD averages for science [7], [8]. Many possible factors may influence students’ scientific literacy, such as educational facilities, curriculum content, programs, learning activities, and social implications. Gender, cultural, economic, and social factors have been directly contributing. The gender difference is one issue that has been widely covered and is also included in science-related studies and its relationship to academic achievements in science [9], [10]. Gender difference is a state of affairs in which there is likely to be a substantial gap in achievement between boys and girls [9], [11].

Science is considered to be a part of 21st-century culture. Science is well structured to provide opportunities for critical thinking, the use of factual data, to learn the values of controlled and comparative research, and to expand the ability to solve new problems [12]. Learning science is about accomplishing science concepts along with the ability to apply these concepts and theories to real-life situations to solve simple or complex issues people find every day [13]–[15].

In scientific literacy researches, the majority of data analysis used Microsoft Excel software and parametric and non-parametric procedures in the IBM Statistical Package for Social Sciences (SPSS) Statistics [16]. An alternative to both common approaches is the Rasch Analysis Model (RAM). This program is an ideal solution of data analysis that can be used to analyze responses for educational assessment. RAM is an Item Response Theory (IRT) analysis model developed by Georg Rasch and popularised by Ben Wright using computer software in the form of a Rasch Model System [17]–[20]. One of the benefits about using RAM is that it does not only focus on the raw score when analyzing the ability of the respondents, but it also pays attention to the answer patterns of the respondents, the ability of the respondents to answer the questions, and the difficulty level of the items [11], [21].

From all of the ideas outlined above, this study measures scientific literacy by administering the test in Natural Science Subject. The aim is to analyze students’ achievements in three aspects that compose scientific literacy proposed by the PISA 2015. Scientific literacy includes three types of scientific knowledge, three scientific competencies, and three attitudes towards science. Data are analyzed using the Rasch Analysis Model supported by Winsteps 4.4.5. The results of the study are expected to provide information for the next scientific literacy research as well as for the development of a scientific literacy assessment tool.

2. Methods

This study is a case study which aims to analyze the results of scientific literacy instrument [22]. The participants of this study were 70 ninth-grade students who enrolled in a private junior high school located in West Bandung Regency, West Java, Indonesia. The participants consisted of 40 boys and 30 girls with an average age of 15 years old. Participants are students in their last year and term in junior high school and have learned most of the subject matters in Natural Science Subject.

Data were gathered by administering the instrument in the form of a written test consisting of 30 simple multiple choice questions with four options. PISA 2015 science assessment framework was used as the guide to develop the instrument. All of the scientific literacy aspects, such as scientific knowledge, scientific competencies, and attitudes toward science, were measured using the same instrument. The questions also covered individual, national/local, and international contexts in various themes within the natural science scope for junior high school.
After collecting the data, it was then analyzed using Rasch Analysis Model (RAM). RAM was chosen because of the benefits which had been elaborated previously. To support the analysis, RAM was assisted by Winsteps 4.4.5. To identify the scientific literacy of students, item reliability, person reliability, variable (Wright) map, and differential item functioning (DIF) were analyzed. Variable (Wright) map provided information regarding students’ abilities and the difficulty level of questions or items. DIF was explored to discover if there was gender bias caused by the instrument. In the analysis, the respondents and gender were coded. For example, code 08B means respondent number 08, and the gender is a boy. For girls, the letter after numbers is changed to b, which stands for the girl.

3. Result and Discussion

3.1. person and item reliability

By analyzing data using Rasch Analysis Model (RAM), the results provide three kinds of reliability values. These are overall instrument reliability shown by the value of Cronbach Alpha, person reliability shown by person-reliability value, and item reliability shown by item-reliability value. The item-reliability value indicates the quality of questions, while person-reliability gives information about the consistency of student’s answers [11], [16]. The following Figure 1 presents the summary statistics of analysis, which includes all three reliability values.

![Figure 1. Person-reliability analysis result](image)

Based on Figure 1, the overall reliability shown by Cronbach Alpha (KR-20) value of .50 means that the interaction between person and item is in bad category because it is in range of .50 - .60. The person-reliability value is .47, which means it also falls into the weak category because it is below .67. Person measure can give information about the average score of respondents in the instrument [15]. The person-measure value is .09 and is very close to the default set at .0. This value can be interpreted that there is no tendency of overall respondents answering correct or incorrect questions in the instrument. For additional information, the average value, which is higher (positive value) or lower (negative value) than 0.0 shows that there is a tendency of respondents’ answers. In this case, the average ability of respondents or students is at the same level as the item difficulty.

Another data that can be gained from Figure 1 is the INFIT MNSQ and OUTFIT MNSQ. These values can tell about the respondents’ suitability in this research. Infit and Outfit Mean Square (MNSQ) values are .99 and 1.05, respectively. These are close to the ideal value of 1.00. The values mean that the students are very suitable to be respondents in this research.

The grouping of respondents can be known by separation value. By the value of .94, the instrument has a very good quality because it can identify the groups of the respondents. Another way to find the groups in such detail manner is by using the equation \( H = [(4 \times SEPARATION) + 1]/3 \). Thus, with
the person separation value of .94, \( H = [(4 \times 0.94) + 1]/3 \) equals 1.59 and is rounded up to 2. Therefore, there are two groups of the respondent in term of ability.

![Summary of 30 Measured Item](image)

**Figure 2.** Item-reliability analysis result

Figure 2 shows that the reliability of item is in very good category because the value of .95 falls into the highest category which has the range value more than .94. Thus, from the person and item reliability values (.47 and .95 respectively), the answer consistency of respondents is weak, however, the quality of items is excellent. The criteria for item suitability are shown by outfit mean square (MNSQ) and outfit (ZSTD) with the acceptable value of outfit MNSQ is in between \( .5 < \text{MNSQ} < 1.5 \) and acceptable value of outfit ZSTD is in between \(-2.0 < \text{ZSTD} < +2.0\). Therefore, the value of 1.05 and .10 for outfit MNSQ and ZSTD respectively show that the items are suitable to be used in this research.

Lastly, the grouping of items in term of difficulty which has already structured in the instrument can be known by the separation value. Using the same equation, the value of 4.23 is input and become \( H = [(4 \times 4.23) + 1]/3 \). The numbers resulted from the calculation is 5.97 and rounded up to 6. This number shows that there are 6 groups or levels of item difficulty.

### 3.2. variable map

Through variable map, the distribution of respondents’ abilities and the distribution of item difficulty will be shown with the same logit scale at once [11]. High logit scores indicate the ability of respondents is high and the increasing item difficulty and vice versa. The map allows the researcher to identify the relationship between the ability of respondents along with the information of their gender and the type of items that they can solve [16]. Both distributions are presented side to side, namely the ability of respondents on the left side and the distribution of item difficulty on the right side as shown by Figure 3.

Based on Figure 3, question number 4 (Q4) is at the highest logit and followed by Q2 and Q26. It means that Q4 is the hardest question in which none can answer correctly. This item is made to measure Interpret Data and Evidence Scientifically competence and Procedural Knowledge with a science proficiency level of 3 in the national context. Even Q2 and Q26 have a different type of competence and level with Q4, which are Evaluate and Design Scientific Inquiries and levels 2 and 5, respectively, three of them share something in common. It is that all of them categorized as the Procedural Knowledge items. Besides, respondents 30G and 40B are the students who got the highest scores. Both highest achievers make it into the same level with question items Q3 and Q9. Question 3 is specified as the item to measure Evaluate and Design Scientific Inquiries and Procedural Knowledge with the level of 3 in the local setting. Meanwhile, Q9 is the attitudes toward science question which is Interest in Science.

On the other hand, Q15 is the easiest question and followed by Q20, Q6, and Q1. All students can answer these four questions correctly. Question number 15 measures Explain Phenomena Scientifically and Epistemic Knowledge with the level of 2 in the personal context. It goes the same for Q1, which
measures the same competence, but it is a Content Knowledge question with the level of 3 in the local context. In contrast, Q20 and Q6 have the same type of attitudes toward science, which is Environmental Awareness. Respondents whose codes 01B and 58G got the lowest scores, and their logits are also equivalent to Q16. This item measures Interpret Data and Evidence Scientifically and Procedural Knowledge with the level of 2 in the personal context.

![Variable (Wright) map presents the distributions of students' ability for scientific literacy (left) and item difficulty (right)](image)

### Figure 3

**3.3. Gender bias**

Gender bias can be detected by Differential Item Functioning (DIF). This analysis is usually used to discover the indication or presence of bias to a certain category of respondents. Bias within items can be identified based on the item probability value which is below 5% or .05 [15]. If an item has a probability value which are less than .05, the item shows bias [11]. The DIF analysis results of this research are presented in Figure 4 as follows.
As shown in Figure 4, there are six item questions whose values are below .05. These are indicated to be having a bias on respondents’ gender. Among six questions that have been indicated of having biases, Q13 and Q14 possessed the highest probability values which are .0348. Both questions have the same level of 3 and science competence of Explain Phenomena Scientifically. In contrast, Q29 has the lowest value of item probability which is .0041. It means this item has an extreme bias on gender. Q29 is an item of attitude toward science which is Interest in Science and Technology.

4. Conclusion

In conclusion, the scientific literacy of junior high school students which is analyzed by Rasch Analysis Model (RAM) falls into the low category. The overall reliability of the instrument is categorised as bad and the person reliability is considered weak. In contrast, the item reliability falls in the very good category which means that the items have excellent quality to measure scientific literacy. The difference of category of person and item reliability shows that respondents’ answer consistency is weak, but the instrument’s items have good quality. From 30 questions, there are six questions which are indicated as having gender biases.

Therefore, scientific literacy needs to have more recognition, especially in science education. Science literacy can be promoted starting from the schools as the main sector of education. Teachers should start to engage students with learning activities that perform scientific literacy indicators. As time goes on, students will be used to materials that have been arranged in scientific literacy manner.

5. References

[1] Garthwaite K, France B and Gillian W 2014 The complexity of scientific literacy: the development
   and use of a data analysis matrix Int. J. Sci. Educ. 36 10 p. 1568–1587
[2] Vieira R M and Vieira C T 2014 Fostering Scientific Literacy and Critical Thinking in Elementary Science Education Int J Sci Math Educ 14 4 p. 659–680
[3] Chanapimuk K, Sawangmek S and Nangngam P 2018 Using science, technology, society, and environment (STSE) approach to improve the scientific literacy of grade 11 students in plant growth and development J. Sci. Learn 2 1 p. 14–20
[4] Voogt J and Roblin N P 2012 A comparative analysis of international frameworks for 21st century competences: Implications for national curriculum policies J. Curric. Stud. 44 3 p. 299–321
[5] Yao J X and Guo Y Y 2018 Core competences and scientific literacy: the recent reform of the school science curriculum in China Int. J. Sci. Educ. 40 15 p. 1913–33
[6] Hastuti P W Setianingsih W and Anjarsari P 2020 How to develop students’ scientific literacy through integration of local wisdom in Yogyakarta on science learning? J. Phys. Conf. Ser 1440 1 p. 012108
[7] OECD 2016 PISA 2015 results in focus (Paris: OECD Publishing)
[8] Purwani L D, Sudargo F and Surakusumah W 2018 Analysis of students’ scientific literacy skills through socioscientific issue’s test on biodiversity topics J. Phys. Conf. Ser 1013 1 p. 012019
[9] Putri A U, Rusyati L and Rochintaniawati D, 2018 The impact of problem-solving model on students’ concept mastery and motivation in learning heat based on gender J. Sci. Learn 1 2 p. 71–76
[10] Aldila M S, Udargo F and Sanjaya Y 2018 Students’ cognitive and attitude of secondary school in learning global warming using student team achievement division (STAD) based on gender J. Sci. Learn. 1 3 p. 104–109
[11] Ringo S S, Samsudin A and Ramalis T R 2020 Utilizing rasch model to analyze a gender gap in students’ cognitive ability on simple harmonic motion J. Phys. Conf. Ser 1467 1 p. 012054
[12] Crowell A and Schunn C 2016 Unpacking the relationship between science education and applied scientific literacy R. Sci. Educ. 46 1 p. 129–140
[13] Holbrook J and Rannikmae M 2009 The meaning of scientific literacy Int. J. Environ. Sci. Educ. 4 3 p. 275–288
[14] Eliyawati, Sunarya Y and Mudzakir A 2017 Solar cell as learning multimedia to improve students’ scientific literacy on science and nanotechnology J. Sci. Learn 1 1 p. 36–43
[15] Juleha S, Nugraha I and Feranie S 2018 The effect of project in problem-based learning on students’ scientific and information literacy in learning human excretory system J. Sci. Learn 2 2 p. 33–41
[16] Suryana T G S, Setyadin A H and Kaniawati I 2020 Assessing multidimensional energy literacy of high school students: an analysis of rasch model J. Phys. Conf. Ser 1467 1 p. 012034
[17] Rasch G 1966 An item analysis which takes individual differences into account Br. J. Math. Stat. Psychol. 19 p. 48–57
[18] Olsen L W 2003 Essays on georg rasch and his contributions to statistics (Unpublished PhD Thesis at Institute of Economics University of Copenhagen)
[19] Bond T G and Fox C 2015 Applying the rasch model fundamental measurement in the human sciences 3rd ed. (New York: Routledge)
[20] Sumintono B 2018 Rasch model measurements as tools in assessment for learning in International Conference on Education Innovation 1 2 p. 71-76
[21] Andriani N, Suhendi E, Samsudin A and Ramalis T R 2020 Identification of creative thinking ability of malay ethnic students in archimedes law with rasch analysis model (RAM): a case study J. Phys. Conf. Ser 1467 1 p. 012050
[22] Mertler C A 2017 Improving schools and empowering educators 5th ed. (California: SAGE Publications)