ITEM CHARACTERISTICS ANALYSIS USING THE RASCH MODEL IN THE DEVELOPMENT OF READING LITERACY INSTRUMENTS FOR ELEMENTARY SCHOOLS STUDENTS

I Nyoman Indhi Wiradika[1], Heri Retnawati[2]
[1]Informatics Education, Ganesha University of Education, Indonesia
[2]Postgraduate Program, Yogyakarta State University, Indonesia

ABSTRACT

Literacy is a basic reading skill in modern society. The development of reading literacy questions is needed to be useful as an assessment function. The development of reading literacy questions needs to pay attention to the principle of conformity with indicators and pay attention to the characteristics of the items. A total of 167 fifth grade elementary school students were sampled in the test instrument for the reading literacy test. Instrument testing was carried out in two stages; first, content validity testing by seven experts, secondly analysis of model fit and item difficulty to determine item characteristics. The results showed that based on the results of expert judgment, 17 items were valid and 1 item had moderate validity. Items with moderate validity are S18 items that are classified as evaluating and reflecting type questions at level 6. Further analysis using the Rasch model obtained 2 items in the very easy category, 14 items in the easy category, 1 item in the difficult category and 1 item in the difficult category. very difficult category. With a distribution of difficulty levels, 11% of very easy questions, 78% of easy questions, 6% of difficult questions, and 6% of very difficult questions.

Corresponding Author:

I Nyoman Indhi Wiradika,
Informatics Education, Faculty of Engineering and Vocational,
Ganesha University of Education,
Jalan Udayana No 11, 81116, Singaraja, Bali, Indonesia
Email: iwiradika@undiksha.ac.id

1. INTRODUCTION

It is expected that reading literacy is often used as a fundamental prerequisite competency for success in school and is believed to be a fundamental skill in modern society. Rintaningrum's research results (2019) revealed that reading literacy has many benefits for readers such as religious reasons, self-development, professional development, problem-solving, personal branding enhancement, international participation, school success, family pride, helping teachers, helping other students, self-confidence, acquire more language knowledge, syntax, visualization skills, recognize reading patterns, improve reading skills, improve writing skills, develop knowledge and information, references, entertainment, be calm and wise, and increase concentration power. Therefore, it is undeniable that reading skills increase people's insight and that information helps them become agents of national change (Rintaningrum, 2019).

There are many views regarding the definition of literacy. Literacy is popularly defined as reading and writing, which is necessary for communication. However, literacy is more than just reading and writing, and it is indeed a term commonly used by researchers (Naidoo, Reddy, & Dorasamy, 2014). The meaning of literacy has expanded in recent years beyond the functional understanding of literacy competencies to include visual and digital literacy in conjunction with written word comprehension (Murphy, Conway, Murphy, & Hall, 2014).

The PIRLS definition of reading literacy is based on a 1991 IEA study, where reading literacy was defined as "the ability to understand and use forms of written language required by society and/or valued by individuals (Mullis & Martin, 2019). With successive assessments Successively, this definition has been elaborated so as to retain its applicability to readers of all ages and forms of written language but makes explicit reference to aspects of the reading experience of young students as they become advanced readers, highlighting the widespread importance of reading in school and everyday life, and acknowledging the increasing variety of texts in today's technological world. Currently, the definition of reading literacy according to PIRLS is the ability to understand and use the forms of written language required by society and/or valued by individuals. Readers can construct meaning from texts in various forms. They read to be to learn, to participate in the community of readers in school and everyday life, and for enjoyment (Mullis & Martin, 2019).

The need for literacy skills needs to be a concern in education. Literacy results in Indonesia still find several problems. PISA, through the OECD, conducts an assessment related to literacy in children aged 15 years. In 2018 Indonesia was below the average reading score of PISA participants. In 2018 there was an increase in the literacy rate between men and women. This means that there is an increase in female illiteracy, which means the gender gap in education between men and women is increasing (Harahap, Maipita, & Rahmadana, 2020).

The Ministry of Education, Culture, Research and Technology of the Republic of Indonesia is trying to overcome the Literacy Problems by issuing a policy of the School Literacy Movement and Minimum Competency Assessment (AKM). The implementation of the School Literacy Movement Program is carried out in stages by considering the readiness of schools throughout Indonesia. This readiness includes the readiness of school capacity, the readiness of school residents, and the readiness of other support systems (Teguh, 2020). The Minimum Competency Assessment (AKM) is an assessment of the minimum ability that includes reading literacy and numeracy by students at a certain level. The policy change towards literacy and numeracy learning encourages schools to implement reading and numeracy literacy learning practices. Many education units do not understand well the implementation of national assessments, so teachers have not prepared anything (Rokhim et al., 2021).

The results of the reading culture evaluation report conducted by Sulistyo (2017) show that the level of success in five literacy learning activities in one school is still at the level of student involvement, including reading 15 minutes before learning, summarizing, writing diaries, mass reading and making wall magazines, while the other two are already at the quality level, namely poetry reading and speech reading activities.

Policies related to literacy in Indonesia have been developed through the school literacy movement. The development of reading literacy is not only related to activity-based programs. Over the past year, the Indonesian government has focused on developing reading literacy as a component of the national assessment. The big change in orientation occurred because it succeeded in replacing the national exam with a standardized test that had been running for decades.

The development of the reading literacy test refers to the literacy framework and cognitive processes. Specifically for reading competence, the sub scale used is students' ability to obtain information (retrieving information), interpret text (interpreting text), and reflect on text (reflecting text) (Musfiroh & Listyorini, 2016).

The National Assessment Assessment includes three aspects, namely the Minimum Competency Assessment (AKM), the character survey, and the learning environment survey (Novita, 2021). The Minimum Competency Assessment (AKM) is an assessment of the basic competencies needed by all students in order to develop their own abilities and play an active role in society in activities that have positive values. AKM is used to measure students' cognitive abilities where the aspects measured are reading literacy and numeracy literacy. AKM is designed to encourage the implementation of innovative learning that is oriented to the development of reasoning abilities, not focusing on memorization. At the same time, the character survey was carried out to measure the mastery of the Pancasila principles by students and their implementation. Thus, it is hoped that a conducive learning environment will be created.

Table 1. Components of Literacy AKM

Component	Competence
Content	Information Text and Literacy Text
Context	Personal, Socio-Cultural, and Scientific
Cognitive Process	Finding Information, Understanding, Evaluating and Reflecting

Reading is a multidimensional domain. While many elements are part of the construction, not all can be taken into account in constructing a Literacy assessment. In PISA, only a few aspects that are considered the most important are selected. The PISA reading literacy assessment builds on three main task characteristics to ensure broad domain coverage:

1. process, which refers to a cognitive approach that determines how readers engage with a text
2. text, which refers to the scope of the material read
3. situation, which refers to the broad range of contexts or purposes in which the reading takes place

Table 2. Grouping literacy levels

Level	Descriptor
1	Learners have the language skills to communicate and reason in accordance with the objectives to peers and adults about themselves and the surrounding environment. Students are able to understand and convey messages; express feelings and ideas; participate in conversations and discussions politely. Students are able to improve their mastery of new vocabulary through various language and literary activities with various topics.
2	Students are able to understand and convey ideas from informational texts, understand characterizations and messages from narrative texts. Students are able to express ideas in group work and discussion. Students are able to improve their mastery of new vocabulary through various language and literary activities with various topics. Students are able to read fluently.
3	Students are able to understand, process, and interpret information and messages from oral and written presentations on topics recognized in narrative and informational texts. Students are able to respond and present the information presented; actively participate in discussions; write down his response to the reading using his experience and knowledge; write texts to convey their observations and experiences in a more structured way. Students have the habit of reading for entertainment, increasing knowledge, and skills.
4	Students are able to understand, process, and interpret exposure information on various topics and literary works. Students are able to actively participate in discussions, present, and respond to non-fiction and fiction information presented; Students write various texts to convey their observations and experiences in a more structured manner, and write their responses to exposures and readings using their experiences and knowledge. Students develop self-competence through exposure to various character strengthening texts.
5	Students are able to understand, process, interpret, and evaluate information from various types of texts on various topics. Students are able to synthesize ideas and opinions from various sources. Students are able to actively participate in discussions and debates. Students are able to write various texts to express opinions and present and respond to non-fiction and fiction information critically and ethically.
6	Students have the language skills to communicate and reason in accordance with the goals, social, academic, and work contexts. Students are able to understand, process, interpret, and evaluate various types of texts on various topics. Students are able to create ideas and opinions for various purposes. Students are able to actively participate in language activities that involve many people. Students are able to write various texts to reflect and actualize themselves to always work by prioritizing the use of Indonesian in various media to advance the nation's civilization.

In the PISA assessment, text features and process variables (but not situation variables) are manipulated to influence task difficulty. The process is manipulated through the goals set in the task.

The levels on PISA provide a useful way to explore the development of reading literacy demands on the composite scale and each sub-scale. The scale summarizes both a person's proficiency in terms of his abilities and the complexity of an item in terms of difficulty. Mapping students and items on one scale represent the idea that students are more likely to successfully complete a task that is mapped at the same level on the scale (or lower) and less likely to be able to successfully complete a task that is mapped on the same scale.

The level-based assessment approach has important information for mapping learning outcomes. The results obtained on the assessment by level provide useful information to support the learning process. Teaching at the Right Level (TaRL) is an evidence-supported educational approach that helps children develop basic

reading and math skills. The TaRL program is getting more attention because of its high effectiveness in improving learning outcomes (Muralidharan, Singh, & Ganimian, 2019). The mapping of literacy competencies is shown at Table 2.

2.   RESEARCH METHOD

This research is development research by developing a reading literacy instrument for elementary students. The stages of instrument development were carried out by developing a reading literacy instrument framework, performing face validation for seven teachers, analyzing with Rasch to determine the level of difficulty and item characteristics. The parameters of persons and items in the Rasch model are expressed on a logit scale. Given the fundamental equation of the Rasch model using a logistic function, continuous data such as the total score (the sum of item scores) can be converted into an interval scale (Putra & Retnawati, 2020).

The research was conducted in Mataram City, West Nusa Tenggara. The reason for choosing the city of Mataram as a sample is because Mataram City is the target of using a computer-based Literacy assessment in 2023. In addition, the city of Mataram also has several programs that aim to improve students' literacy competence, thus requiring quality literacy instruments.

The sample in this study was grade 5 elementary school students. The selection of grade 5 students is based on a minimum competency assessment policy that measures children's literacy skills since grade 5. The number of samples in this study was 167 students, with 99 female students and 68 male students. The sample selection was carried out representatively representing the sub-districts in Mataram City.

3.   RESULT AND DISCUSSION

The validity test in this study uses content validity obtained through expert judgment. In this study, experts or teachers were asked for their opinions regarding the suitability between question items in terms of content and editorial items. Furthermore, expert test considerations are used as a basis for researchers to improve instrument items.

Content validation in this study was carried out by expert judgment or experts, namely seven fifth grade elementary school teachers. The criteria in determining item decisions are seen based on the index, if it is less or equal to 0.4 then the validity is less, while 0.4-0.8 is moderate validity, if it is more significant than 0.8 it is said to be valid (Retnawati, 2016). Following are the results of the validity of each variable with the Aiken formula:

Table 3. Content validity of Literacy

Question Code	Aiken Score	Category	Question Code	Aiken Score	Category
S1	0.0625	valid	S10	0.0597222	valid
S2	0.0659722	valid	S11	0.05625	valid
S3	0.0597222	valid	S12	0.05625	valid
S4	01.00	valid	S13	0.0625	valid
S5	0.0659722	valid	S14	0.0625	valid
S6	0.0659722	valid	S15	0.0597222	valid
S7	0.0597222	valid	S16	0.05625	valid
S8	0.0659722	valid	S17	0.0597222	valid
S9	0.0659722	valid	S18	0.0465278	moderate validity

Table 3 shows 17 valid items and 1 item with moderate validity. Items with moderate validity are S18 items classified as evaluating and reflecting type questions at level 6. Because there are no items that fall into the category of less validity, all items were used in trials for fifth-grade elementary school students and analyzed using the Rasch model.

The analysis with the Rasch model produces a statistical analysis of suitability (fit statistics) which provides information to the researcher whether the data obtained ideally illustrates that people who have high abilities provide patterns of answers to items according to their level of difficulty.

In the first stage, the overall fit model was tested for applying the Rasch model. The method used is Andersen LR (Likelihood ratio) test (Andersen, 1973). If the value is not significant ($p > 0.05$), it indicates that the data fit the model. The model fit test results. Look at the p-value, and it turns out that the p-value = 0.24 is greater than the standard (0.05), meaning the model is good enough (good fit of the model).

The results of the analysis with the R program through the 'ltm' package for the analysis of the Rasch model obtained the difficulty level of each question using equation 1 in the R programmer, the results of the difference in power of questions (a) are worth one and the difficulty level of questions per item (b) is shown in Table 4

Table 4. The level of difficulty of the questions

Question Code	Difficulty (b)	Category	Question Code	Difficulty (b)	Category
S1	-3.41505	very easy	S10	-0.02308204	easy
S2	-1.78367	easy	S11	0.32134815	easy
S3	-3.56333	very easy	S12	0.12023268	easy
S4	-0.45988	easy	S13	-0.86453454	easy
S5	-1.06647	easy	S14	-0.10878326	easy
S6	-1.20893	easy	S15	-0.13792097	easy
S7	-1.51934	easy	S16	0.6149696	easy
S8	-2.32694	easy	S17	0.79787444	difficult
S9	-2.2643	easy	S18	2.56861391	very difficult

By referring to the index of the difficulty level of each item in Table 4 which has the lowest value of -3.563 and the highest of 2.568, the average difficulty level (brata2) is -0.795 with a standard deviation (bstdev) of 1.540. The categorization of questions, according to Sumintono & Widhiarso (2015), refers to the criteria contained in Table 5 below:

Table 5. Categorization of item difficulty levels

| Criteria	Item Difficulty (b)	Category
b <(bmean- bstdev)	-2.336	very easy
(bmean - bstdev) <b <bmean	-2.336 <b <-0.79551	easy
bmean <b <bstdev	TRUE	difficulty
b >bstdev	1.54005	very difficult
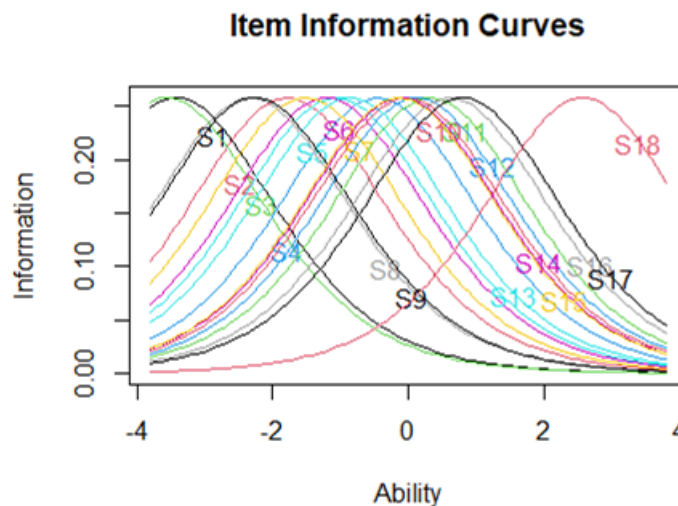
Figure 1. Graph of the measurement information function (test information function)

So, it was found that two items were in the very easy category, 14 items were included in the easy category, 1 item was in the difficult category and 1 item was included in the very difficult category. With a distribution of difficulty levels, 11% of very easy questions, 78% of easy questions, 6% of difficult questions, and 6% of very difficult questions. So, the reading literacy question cannot be used to measure students with various literacy abilities. The questions that fall into the very easy category and are in the student's ability scale range from -3.454 to -3.605 are questions about finding information and evaluating simple texts. The questions categorized as difficult and very difficult are questions with a student's ability scale from 0.827 to 2.622 to determine students' literacy skills about complex inference and evaluation of complex texts.

Each measurement should present information about the measurement results. In this case, the information that is measured is not information about the individual respondents but information about the items, where the information is related to the questions and the results of student responses to the questions (Mardapi, 2016). The information obtained is very dependent on the variation in the measurement results, namely student responses. So that more respondents will provide actual variations in the results of measuring items.

The X-axis in Figure 1 shows the level of students' abilities which are made on a scale from very low (ability or ability -4), low (ability scale or ability -2), medium (ability 0), high (ability 2) and very high (abilities 4). The Y axis shows the magnitude of the information function obtained from the results of the 16 item questions. The following will analyze three graphs of the item information function that represent the group of questions for items no. S1, S3, S17, and S18 with very easy, difficult and very difficult categories, as shown in Figures 2(a), 2(b), and 3.

| (a) | (b) |
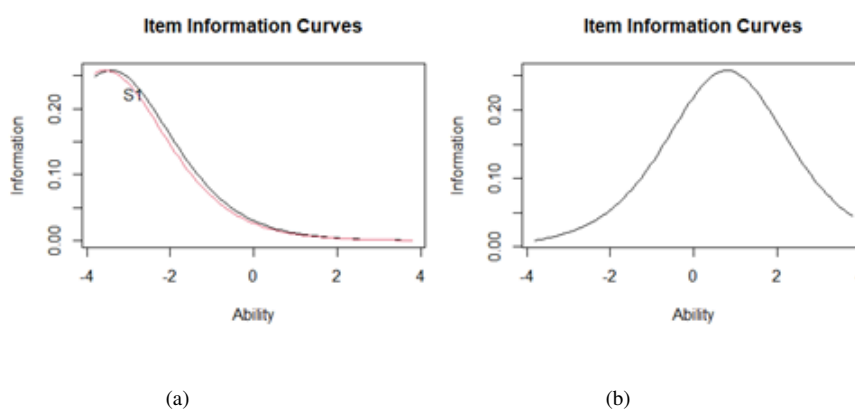
Figure 2. Graph of the information function,
(a) S1 and S3 items with very easy categories, (b) the 17th item in the difficult category
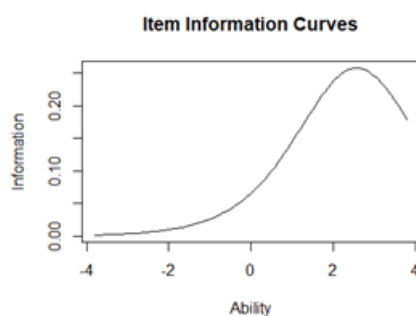
Figure 3. The graph of the information function of the 18th item with a very difficult category

The function of information in showing the value of reliability is independent against others. This is because the information function is part of the analysis results using item response theory which can correct the shortcomings of the classical test theory.

The items for S1 (Figure 3a) and S3 (Figure 3b) are level 1 reading literacy questions. The S1 questions aim to test students' ability to find and seek information. S3 questions aim to test students' ability to evaluate and reflect on texts.

The S1 test item has a parameter of b = -3.415, which means that to solve this problem, it is necessary to have the ability of students on a minimum scale of -3.451 to be able to answer correctly with a 50% chance.

While the S3 test items require students' abilities on a minimum scale of -3.563 to be able to answer correctly with a 50% chance.

Figure 4. Item question,
(a) S1, (b) S3

Items S17 (Figure 4b) and S18 (Figure 3c) are level 6 reading literacy questions. The S17 questions aim to test students' ability to infer text. The S18 question tests students' ability to evaluate and reflect on complex texts. Items S17 and S18 are designed to have a high degree of difficulty to ascertain the level of students' abilities. However, based on the results of expert validity, it was found that item S18 was included in moderate validity, and the results with the Rasch model showed that item S18 was very difficult.

Figure 5. Question Text S16, S17 and S18

17. Berdasar teks di atas, menurutMU apa yang menjadi masalah bagi kaum difabel di tanah air saat ini?
A. Persepsi masyarakat tentang difabel dan minimnya fasilitas layanan pub difabel
B. Sering menjadi korban Hak Asasi Manusia seperti penculikan terhadap p disabilitas
C. Minimnya keberadaan kursi roda yang canggih untuk kebutuhan sehari-
D. Penampilan dan tigkah laku yang tidak sama seperti orang-orang pada u
E. Dokumen-dokumen ilmiah tentang difabel jenis-jenis disabilitas sulit dit......

18. Mengapa penulis mengutip Undang-Undang (UU) Nomor 8 Tahun 2016 dalam teks?
A. Mengatur tentang fasilitas yang wajib diberikan kepada difabel
B. Aturan yang menjadi dasar pengelolaan bantuan kepada difabel oleh Kemensos
C. Mengatur HAM tentang hak warga Negara untuk memperoleh pendidikan
D. Salah satu dasar regulasi Menteri dalam menentukan pejabat
E. Standar penentuan Lencana Karya Satya dari Presiden RI

(a)   (b)

Figure 6. Question Text,
(a) S17, (b) S1

The item S17 test has a parameter of $b = 0.797$, which means that to solve this problem, students' abilities on a minimum scale of 0.797 are required to answer correctly with a 50% chance. While the S18 test items require students' abilities on a minimum scale of 2,568 to be able to answer correctly with a 50% chance.

From the overall analysis of this item, it was found that by analyzing the information function of the item, it was very helpful to show the measurement function of the item and analyzing the probability of students answering correctly with certain abilities in the item.

4.   CONCLUSION

Based on the results of expert judgment obtained 17 valid items and 1 item with moderate validity. Items with moderate validity are S18 items classified as evaluating and reflecting type questions at level 6. Further analysis using the Rasch model obtained 2 items in the very easy category, 14 items in the easy category, 1 item in the difficult category and 1 item in the difficult category. very difficult category. With a distribution of difficulty levels, 11% of very easy questions, 78% of easy questions, 6% of difficult questions, and 6% of very difficult questions.

The findings of this study resulted in as many as 17 items that could be used to measure students' literacy skills. Item S18 needs improvement because it is too difficult for the context of fifth grade elementary school questions. Improvements to the S18 items need to be carried out by considering expert opinions and need to be re-analyzed with the Rasch model.

REFERENCES

Andersen, E. B. (1973). A goodness of fit test for the Rasch model. *Psychometrika*, *38*(1), 123–140.

Harahap, E. S., Maipita, I., & Rahmadana, M. F. (2020). Determinant Analysis of Education Inequalities in Indonesia. *Budapest International Research and Critics Institute (BIRCI-Journal): Humanities and Social Sciences*, *3*(2), 1067–1082. doi:10.33258/birci.v3i2.937

Mardapi, D. (2016). *Pengukuran, Penilaian dan Evaluasi Pendidikan*. Yogyakarta: Parama Publishing.

Mullis, I. V., & Martin, M. O. (2019). PIRLS 2021 Reading Assessment Framework. Retrieved from http://pirls2021.org/frameworks/home/reading-assessment-framework

Muralidharan, K., Singh, A., & Ganimian, A. J. (2019). Disrupting Education? Experimental Evidence on Technology-Aided Instruction in India. *American Economic Review*, *109*(4), 1426–1460. doi:10.1257/aer.20171112

Murphy, B., Conway, P. F., Murphy, R., & Hall, K. (2014). The emergence of reading literacy in post-primary teacher education: from the background to the foreground. *European Journal of Teacher Education*, *37*(3), 331–347. doi:10.1080/02619768.2013.870995

Musfiroh, T., & Listyorini, B. (2016). KONSTRUK KOMPETENSI LITERASI UNTUK SISWA SEKOLAH DASAR. *LITERA*, *15*(1). doi:10.21831/ltr.v15i1.9751

Naidoo, U., Reddy, K., & Dorasamy, N. (2014). Reading Literacy in Primary Schools in South Africa: Educator Perspectives on Factors Affecting Reading Literacy and Strategies for Improvement. *International Journal of Educational Sciences*, *7*(1), 155–167. doi:10.1080/09751122.2014.11890179

Putra, M. D. K., & Retnawati, H. [Heri]. (2020). Rasch Analysis of an Indonesian version of the General Self-Efficacy Scale-12: A Comparison of Rating Scale Model (RSM) and Partial Credit Model (PCM). *Indonesian Journal of Educational Assesment*, *3*(1), 52. doi:10.26499/ijea.v3i1.55

Retnawati, H. [H.]. (2016). *Validitas Reliabilitas dan Karakteristik Butir*. Yogyakarta: Nuha Medika.

Rintaningrum, R. (2019). Explaining the important contribution of reading literacy to the country's generations: Indonesian's perspectives. *International Journal of Innovation, Creativity and Change*, *5*(3), 936–953.

Rokhim, D. A., Rahayu, B. N., Alfiah, L. N., Peni, R., Wahyudi, B., Wahyudi, A., . . . Widarti, H. R. (2021). ANALISIS KESIAPAN PESERTA DIDIK DAN GURU PADA ASESMEN NASIONAL (ASESMEN KOMPETENSI MINIMUM, SURVEY KARAKTER, DAN SURVEY LINGKUNGAN BELAJAR. *Jurnal Administrasi dan Manajemen Pendidikan*, *4*(1), 61. doi:10.17977/um027v4i12021p61

Sulistyo, A. (2017). EVALUASI PROGRAM BUDAYA MEMBACA DI SEKOLAH DASAR NEGERI. *Kelola: Jurnal Manajemen Pendidikan*, *4*(1), 48. doi:10.24246/j.jk.2017.v4.i1.p48-58

Sumintono, B., & Widhiarso, W. (2015). *Aplikasi Pemodelan Rasch Pada Assessment Pendidikan*. Cimahi: Trim Komunikata.

Teguh, M. (2020). GERAKAN LITERASI SEKOLAH DASAR. *Jurnal Pendidikan Dasar Flobamorata*, *1*(2), 1–9.