

## Pirls-based cognitive assessment instrument: How effective is it for measuring Indonesian language reading literacy in fifth grade?

Yuli Putri Pratiwi\*  
Nurlaksana Eko Rusminto  
Ryzal Perdana

Faculty of Teacher Training and Education, Universitas Lampung, Bandar Lampung, Indonesia

### ABSTRACT

The Indonesian reading literacy test is a technique for evaluating students' cognitive reading abilities. However, as of yet, the test questions administered to students are inadequate to assess certain parts of their Indonesian reading literacy. Therefore, the purpose of this study was to develop an adequate and standardized cognitive assessment instrument based on the PIRLS model to measure students' Indonesian reading literacy. The participants in the study were fifth graders from an elementary school, SD Negeri 11, located in Metro City, Lampung Province, Indonesia, with a total of 293 students. This study adopted a research and development design. The results indicated that the cognitive assessment instrument based on the PIRLS paradigm that was developed to measure Indonesian reading literacy in fifth graders delivered adequate and standardized items.

### KEYWORDS

Cognitive assessment instruments; PIRLS model; reading literacy.

Received: 24 July 2022  
Accepted: 19 August 2022  
Published: 20 August 2022

### Introduction

Students' reading literacy is the main factor in digesting learning because most of the knowledge is presented in written language (Yohana et al., 2019). Literacy is the ability to identify, understand, translate, create, communicate and process the contents of a series of texts in printed and written materials related to various situations (Aisyah et al., 2017). Literacy is categorized into three forms of text that are most often encountered by students at school and in everyday life, namely: literacy examines three main things, namely: (1) narrative prose, (2) expository prose, and (3) documents (Muhammadi et al., 2018). The latest understanding of the meaning of literacy includes the ability to read, understand, and appreciate various forms of critical communication, which includes spoken language, written communication, and communication that occurs through printed or electronic media (Wardana & Zamzam, 2014).

In the context of learning, literacy plays an essential role in achieving learning objectives both at the micro and macro levels (Tryanasari, 2017). Thus, unavoidably, students must have the skills and willingness to read to increase their knowledge. Therefore, reading literacy in elementary schools is the foundation or basis for determining the success of students learning at the next level.

Reading is an important and essential skill for the education system, which is a significant part and tool of the world's goal of seeking knowledge, generating ideas, and making decisions and solutions (Satthapong, 2018). Reading is an activity to gain meaning, information, knowledge and experience from printed/written media, books, magazines, or other writings.

Elementary school education level (SD) is a level of education that is very decisive in the success of students seeking knowledge to move to a higher level. Elementary school is a continuous learning process from playgroups, PAUD, and kindergartens. In the school environment, learning processes are introduced which are not only about real life but also theories and the process of being educated for the better for changes in ideas and behavior. Especially in elementary school, it is divided into low class and high class. The lower class includes students in grades I, II, and III. While the high class includes IV, V, and VI grades (Prime et al., 2021).

CONTACT Yuli Putri Pratiwi  [yuliputripratiwi26@gmail.com](mailto:yuliputripratiwi26@gmail.com)

© 2022 The Author(s). Published with license by Lighthouse Publishing.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

The reading literacy ability of elementary school students in Indonesia can be seen compared to several countries worldwide. The Progress in International Reading Literacy Study (PIRLS) in 2006 reported that the reading ability of fourth-grade elementary school students in Indonesia was still relatively low. In contrast, the sixth-grade Indonesian elementary school students' reading proficiency with a score of 51.7 was below the Philippines (52.6), Thailand (65.1), Singapore (74.0) and Hong Kong (75.5), (Gumono, 2013). The Program for International Students Assessment (PISA) research on the language literacy skills of students from around the world in 2012 showed that Indonesia was ranked 64th out of 65 countries. (Kharizmi, 2015). Furthermore, the EGRA (Early Grade Reading Assessment) study shows that 50% of students in Indonesia can read (literacy), but only half of them understand what is read (Sugiarsih, 2017). Ranking according to World's Most Literate Nations data, compiled by Central Connecticut State University in 2016, Indonesia's literacy ranking is in the second lowest position out of 61 countries studied (Agoestyowati, 2017).

Literacy studies conducted by the International Literacy Institute showed consistent results, namely that the literacy skills of Indonesian children were still low. This means that the reading ability of Indonesian students compared to students in the international world is still inadequate, either at the perfect; high; and medium levels, except at the weak level, which is caused by the lack of seriousness in dealing with reading skills problems, both at the micro and macro education level (Suryaman, 2015).

According to Suryaman (2015), the cause of weak reading and problem solving skills is the tendency of students to answer multiple choice questions by guessing because they are not used to doing literacy test questions. This indicates the weak reading ability of Indonesian children and the unsynchronization of the test with Indonesian conditions. In another words, one of the causes of the low literacy competence of Indonesian children is because they are not accustomed to working on literacy test instruments. The Indonesian reading literacy test instrument is a tool for assessing students' cognitive abilities in reading or the ability to manage and understand information when doing reading activities.

Because of this, the researcher are interested in developing appropriate and standardized cognitive assessment instruments in measuring students' reading literacy skills, which in this case is a literacy test instrument construct adapted from PIRLS. Standardized, meaning that the test questions must go through an analysis of the quality of the test, both as a whole and the items that are part of the test such as question validity, reliability, level of difficulty, discriminatory power and distractors. Feasible, meaning that the test questions are theoretically feasible and practical in use.

This study aimed to develop an appropriate and standardized cognitive assessment instrument based on the PIRLS model to to measure students' Indonesian reading literacy.

## Methods

This type of research was research and development. Research and Development (R&D) is research used to develop products used in education and learning. The research was conducted by following the development research model according to Borg and Gall (Borg & Gall, 1983) which was simplified due to the limitations of the researcher and adapted to research needs into 6 steps including: (1) Research and Information Collecting (2) Planning (3) Develop Preliminary Form of Product (4) Preliminary Field Testing (5) Main Product Revision (6) Main Field Testing.

This research was conducted at SD Negeri 11 Metro, Metro City Center. The subjects of this research were fifth grade students at SD Negeri 11 Metro, Metro City Center in the 2021/2022 academic year. In addition, the object of this research was the development of a cognitive assessment instrument based on the PIRLS model to measure standardized Indonesian reading literacy in class V SD Negeri 11 Metro, Metro City Center.

Data collection techniques were carried out through documentation, questionnaires and tests using questionnaires and tests. In order to determine the feasibility of the product that was developed theoretically, it was carried out through expert testing, while in terms of practicality, a questionnaire response from educators and students was used. In order to produce a standardized cognitive assessment instrument, it was carried out through validity testing, reliability testing, difficulty level testing, and discrimination power tests.

## Results

### Validity Test

The distribution of items based on the the category of question validity can be seen in the following table.

**Table 1.** Distribution of Question Item Based on Validity

Criteria	Question Number	Total	Percentage
Valid	1,2,3,5,6,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,25,26,27,28,29,30	26	86,67%
Invalid	4,7,12,24	4	13,33%

Table 1 shows that the distribution of validity is as many as 26 or 86.67% of the items are valid while the remaining 4 or 13.33% items are invalid.

### Reliability Test

A research instrument is said to have a high-reliability value if the test made has consistent results in measuring what is to be measured.

**Table 2.** Result of Reliability Test

Cronbach's Alpha	N of Items
.844	26

Based on the test results, the reliability value of the item was 0.844. This shows that the items tested are reliable in the very high category (Riduwan, 2009).

### Difficulty Level Test

The distribution of items based on the level of difficulty can be seen in the following table.

**Table 3.** Distribution of Question Item Based on Difficulty Level

Category	Question Number	Total	Percentage
Easy	2,6,10,12,18,19,23,25	8	30,8%
Medium	1,4,5,9,11,13,14,15,17,20,21,22,24,26	14	53,8%
Difficult	3,7,8,16	4	15,4%

Table 3 shows that the distribution of items based on the difficulty level is as many as 8 or 30.8% items are in the easy category, as many as 14 or 53.8% items are in the medium category. As many as 4 or 15.4% items are in the difficult category.

### Discrimination Power Test

The distribution of items based on their discriminating ability can be seen in the following table.

**Table 4.** Distribution of Question Item Based on Determinant Power

Category	Question Number	Total	Percentage
Bad ( <i>poor</i> )	6	1	3,8%
Fair ( <i>satisfactory</i> )	2,7,12,13,14,17,19,22,23,25	10	38,5%
Good ( <i>good</i> )	1,3,4,5,8,9,10,11,15,16,18,20,21,24,26	15	57,7%
Very Good ( <i>excellent</i> )	-	0	0%

Table 4 shows that the distribution of items based on their discrimination power is 1 or 3.8% of items with poor discriminating power, as many as 10 or 38.5% of items have sufficient discriminating power, and as many as 15 or 57.7% of the items have good discriminating power.

### Distractor Effectiveness Test

Based on the results of the distractor effectiveness test of the items, all 25 items have good distractors where all distractors have a percentage of more than 5%.

## Discussion

### Validity Test

Based on the results of the validity test, of the 30 items tested, 4 questions were declared invalid because they had an r-count which was smaller than r-table. As for the invalid items, they were questions number 4, 7, 12 and 24, which means they are items that cannot measure the expected competence and henceforth will no longer be used. While the other 26 items, were declared valid because these questions could measure the expected competence.

Research related to the analysis of the content validity of the questions was also carried out by Orihabor & Emafo (2016). Their research entitled "Determining the Reliability and Content Validity of the Mathematical Tests Constructed by Senior Secondary School Mathematical Teachers in Edo State, Nigeria" indicates that the questions made by the teacher have sufficient internal reliability and low content validity.

To find out whether the tests (questions) used have met the criteria for a good test, it is necessary to conduct an analysis of the test questions, one of which is by conducting a validity analysis (Utomo, 2018). The results of this study are in accordance with the validity theory according to Sudijono (2012) that items which have high validity reflect that the questions already have reliability and there is no need to doubt their accuracy in measuring students' abilities. For items that have low validity, it reflects that the questions are not valid, thus it is necessary to take action on those questions. The actions taken are either by revising the invalid questions and testing them again or by not using them. In this case, the researcher took the option not to use the questions.

### **Reliability Test**

The procedure for calculating reliability is done by connecting each item in one test with other items in the test itself as a whole. Based on the results of the validity test, 26 valid items were obtained, thus these 26 questions were then tested for reliability. As for the results of the reliability test, it was found that 0.844 was in very high criteria. This shows that the items tested were reliable. This means that these questions had high reliability. The reliability referred to in this case includes the correctness or accuracy of the measurement results and the constancy or stability of the measurement results.

The reliability coefficient shows the extent to which a test consistently places students in the middle of the group. The reliability coefficient is directly influenced by the spread of scores in the measured group (Setiawan, 2014). Therefore, if the test is carried out several times on these test items, it will give the same or relatively the same results.

### **Difficulty Level Test**

The results of the item difficulty level test shows that of the 26 items, 8 or 30.8% of the items are in the easy category. The questions are categorized as easy level because the difficulty index of the questions is in the range of 0.73 - 1.00 (73%-100%). This means that most students can answer the question correctly. The easy category questions are items number 2, 6, 10, 12, 18, 19, 23, 25. Furthermore, there were 14 or 53.8% items in the medium category. The questions are categorized as moderate difficulty because the difficulty index of the questions is in the range of 0.28 - 0.72 (28%-72%). This means that some students can answer the question correctly. The medium category questions are items number 1, 4, 5, 9, 11, 13, 14, 15, 17, 20, 21, 22, 24, 26. In the category of difficult questions, there are 4 or 15.4% of the items in the difficult category. The questions are categorized as difficult because the index of difficulty is in the range of 0.00-0.27 (0%-27%). This means that only a small number of students can answer the question correctly. The questions in the difficult category are items number 3, 7, 8, 16.

In general, an item is declared good if the item is not too difficult and not too easy. Therefore, items that cannot be answered correctly by all students (because they are too difficult) can be declared as bad items. Vice versa, items that all students can answer correctly (because they are too easy) can also be stated as bad items. For both types of categories, improvements need to be made if they will be used again as questions for the next examination (bagiyono, 2017).

Furthermore, Arikunto (2010) explained that a question can be said to be good if the question has a moderate level of difficulty, it means that the question is not too easy or too difficult to do. Easy questions do not make students improve the way they solve them, as well as difficult questions will make students give up and despair because it is beyond their ability. The requirements for the multiple choices test are that the questions consist of 30% easy questions, 50% moderate questions, and 20% difficult questions.

Referring to this opinion, of the 26 questions, 30% of the easy questions are 8 questions, 50% of the medium questions are 13 questions, and 20% of the difficult questions are 5 questions. The results of the study indicated that for easy questions there are 8 or 30.8% questions, medium questions there are 14 or 53.8% questions and in difficult questions there are 4 or 15.4% questions. This means that the proportion of the level of difficulty of the questions has not been fully met. However, the number of each level of difficulty of the questions has approached the ideal.

### **Discrimination Power Test**

The discrimination power of a question is the ability of a question to distinguish between smart students and less intelligent students. A question that can be answered correctly by smart students or less intelligent students, then the question is categorized to be a bad question because it does not have discrimination power. Thus, it can be said that the test tool with good discrimination power has been able to distinguish between better students (high ability) and poorer students (low ability), (Fatimah & Alfath, 2019).

Based on the results of the discriminatory test, it can be seen that as many as 1 or 3.8% of the items have poor discrimination power, namely item number 6. This is due to the possibility that the preparation of the questions does not pay attention to the rules that have been set or the preparation of the questions not according to the rules for making questions (Arisandi et al., 2020). Regarding the items with poor discrimination power, it is recommended that the questions no longer be used.

Furthermore, as many as 10 or 38.5% of the items have sufficient discrimination power, namely items numbered 2, 7, 12, 13, 14, 17, 19, 22, 23, 25. Referring to this, the question can still be used because the questions have discrimination power that can still distinguish between high-ability students and low-ability students. These questions do not have to have high discriminatory power therefore if there are questions with sufficient categories, they can still be used (Magdalena, 2021). In addition, 15 or 57.7% of the items have good discrimination power, namely items number 3, 4, 5, 6, 8, 9, 10, 11, 15, 16, 18, 20, 21, 24, 26.

In terms of discrimination power, Sudijono (2011) suggests that items that already have good enough discrimination power (satisfactory, good and excellent) should be included in questions bank book of the learning outcomes test. Then for items with low discrimination power (poor) they can be revised, replaced or not used.

### ***Distractor Effectiveness Test***

A distractor is a choice of answers or options in the form of multiple choices questions. In those options, there is a correct answer and the other options are called distractors (Rusilowati et al., 2016).

The cognitive assessment instrument developed was in the form of multiple choices questions with a total of 25 questions. Each of questions has 4 options consisting of 1 correct answer option and 3 distractor options. It means that there are 75 distractors on the questions. Of the 75 distractors, all of them have a percentage above 5%, which means that all of the distractors on the questions function well. This is in accordance with the opinion of Sudijono (2011) which states that a distractor has been able to carry out its function properly if the distractor has been chosen by at least 5% of all test participants.

### **Conclusion**

Based on the results of research and discussion, it can be concluded that the cognitive assessment instrument based on the PIRLS model to measure Indonesian reading literacy in class V students that was developed was feasible and standardized both theoretically and practically.

### **References**

- Agoestyowati, Redjeki. (2017). Branding Serial KKP: Review on Interest of Kids Literacy. *Scientific Magazine of Institut STIAM*, (1) 14.
- Arikunto, S. (2010). *Research Procedure: A Practical Approach*. Jakarta: Rineka Cipta.
- Aisyah, D.W., et al. (2017). Pengembangan Bahan Ajar Berbasis Literasi Bercirikan Quantum Teaching untuk Mengoptimalkan Pembelajaran Efektif dan Produktif. *Jurnal Pendidikan: Teori, Penelitian, dan Pengembangan*, 2(5), 667-675.
- Arisandi, Wall Fitri et al. (2020). Analisis Tingkat Kesukaran dan Daya Pembeda Soal Mata Pelajaran Sejarah Kelas VII SMP Negeri 1 Bandar Dua Kabupaten Pidie Jaya. *Educational Journal of History and Humanities*, 3(1) 24-31.
- Bagiyono. (2017). Analisis Tingkat Kesukaran dan Daya Pembeda Butir Soal Ujian Pelatihan Radiografi. *Widyanuklida*, 16(1), 3-14.
- Borg, W.R. & Gall, M.D. Gall. (1983). *Educational Research: An Introduction, Fifth Edition*. New York: Longman
- Fatimah, L. U., & Al-Fath, K. (2019). Analisis Kesukaran Soal, Daya Pembeda, dan Fungsi Distraktor. *Jurnal Komunikasi dan Pendidikan Islam*, 8(2), 41-52.
- Gumono. (2013). Pemanfaatan Bahan Ajar Membaca Berbasis Lokal untuk Meningkatkan Keterampilan Membaca Siswa Kelas IV Sekolah Dasar di Provinsi Bengkulu. *Proceeding International Seminar on Languages and Arts*, 208-209.
- Kharizmi, Muhammad. (2015). Kesulitan Siswa Sekolah Dasar dalam Meningkatkan Kemampuan Literasi. *Jurnal Pendidikan Dasar*, 2(2) 11-21.
- Magdalena, Ina et al. (2021). Analisis Daya Pembeda, Dan Taraf Kesukaran Pada Soal Bilangan Romawi Kelas 4 SDN Tobat 1 Balaraja. *Jurnal Pendidikan dan Ilmu Sosial*, 3(1), 151-158.
- Muhammadi, et al. (2018). Literasi Membaca untuk Memantapkan Nilai Sosial Siswa SD. *Jurnal Litera*, Vol.17 (2): 202-212
- Perdana, Ryzal et al. (2021). Literasi Numerasi dalam Pembelajaran Tematik Siswa Kelas Atas Sekolah Dasar. *Absis: Mathematics Education Journal*, 3(1), 9-15.
- Riduwan, & Akdon. (2010). *Rumus dan Data dalam Analisis Data Statistika*. Bandung: Alfabeta
- Rusilowati, Ani, et al. (2016). Developing an Instrument of Scientific Literacy Assessment on the Cycle Theme. *International Journal of Environmental & Science Education 2016*, 1(12), 518-572.
- Sudijono, Anas. (2011). *Pengantar Evaluasi Pendidikan*. Jakarta: PT Raja Grafindo Persada
- Sugiarsih, S. (2017). Peningkatan Keterampilan Membaca Melalui Drop Everything and Read (Dear) Pada Siswa Sekolah Dasar (MI). *Jurnal Pendidikan Dasar Islam*, 9(2), 47-58.
- Suryaman, Maman. (2015). Analisis Hasil Belajar Peserta Didik Dalam Literasi Membaca Melalui Studi Internasional (PIRLS) 2011. *Jurnal Litera*, 14(1), 170-186.
- Satthapong, Tasanee. (2018). Research and Development of Reading Literacy Follow the Pisa Test Using by Reading Apprenticeship Approach. *The 2018 International Academic Research Conference in Vienna*.
- Setiyawan, Adhi. (2014). Faktor-Faktor yang Mempengaruhi Reliabilitas Tes. *Jurnal An Nûr*, VI(2) 341-354.
- Tryanasari, Dewi., et al. (2017). Pembelajaran Literasi di SDN Rejosari 1 Kecamatan Kawedanan Kabupaten Magetan. *Jurnal Pendidikan Dasar dan Pembelajaran*, 7(2), 173-179.
- Utomo. (2018). Analisis Validitas Isi Butir Soal Sebagai Salah Satu upaya Peningkatan Kualitas Pembelajaran di Madrasah Berbasis Nilai-Nilai Islam. *Jurnal Pendidikan Matematika*, 1(2), 145-158.
- Wardana, & Zamzam. (2014). Strategi Peningkatan Kemampuan Literasi Siswa di Madrasah. *Jurnal Ilmiah Widya Pustaka Pendidikan*. 2(3), 248-258.
- Yohana, et al. (2019). The Development of Instrument of Reading Literacy Assessment on Indonesian Language Learning in Quality Medan University. *Budapest International Research and Critics in Linguistics and Education (BirLE) Journal*, 2(2), 372-387.