Valid and reliable instrument for measuring Indonesian students' reading literacy

Limala Ratni Sri Kharismawati[1], Pratomo Widodo[2], Heri Retnawati[3]
[1]Educational Research and Evaluation, Graduate School, Yogyakarta State University, Yogyakarta, Indonesia
[2]Department of German Language Education, Faculty of Languages, Arts, and Culture, Yogyakarta State University, Yogyakarta, Indonesia
[3]Department of Mathematics Education, Faculty of Mathematics and Natural Science, Yogyakarta State University, Yogyakarta, Indonesia

ABSTRACT

The study aims to verify the validity and reliability of a set of tests for measuring Indonesian students' reading literacy achievement that has been developed by the Center for Language Strategy and Diplomacy Development as well as the Educational Assessment Center, two institutions under the auspices of the Ministry of Education, Culture, Research, and Technology of the Republic of Indonesia. The construct validity of the test was verified using confirmatory factor analysis (CFA) and the reliability was estimated using the composite reliability (CR) formula. The data used were responses from a sample of 6,539 students from 294 high schools in 34 provinces in Indonesia. The sample was established using a multistage sampling technique from the study population of 15-year-old students in Indonesia. The results indicated that the construct validity of the test was verified and the reliability coefficient was 0.805, which was defined good. Based on the results, it can be concluded that the set of tests is appropriate to be used to measure Indonesian students' reading literacy achievement. In other words, using the set of tests will establish measurement that can reflect the reading literacy achievement of Indonesian students.

This is an open access article under the CC BY-SA license.

Corresponding Author:

Limala Ratni Sri Kharismawati
Educational Research and Evaluation, Graduate School, Yogyakarta State University
Campus of Karangmalang, Yogyakarta-55281, Indonesia
Email: limala.ratni2016@student.uny.ac.id, limala.ratnisk@yahoo.de

1. INTRODUCTION

Literacy is no longer interpreted as the ability to read and write. Today, literacy, especially reading literacy, is interpreted as the ability to effectively, efficiently, and simultaneously control various linguistic systems and other sign systems, cognitive, and social-cultural, as well as the development of the dimensions of written language in order to interact [1]. In other words, reading literacy does not only mean as the ability to read written texts. More than that, reading literacy prioritizes the ability of readers to be able to understand and digest texts that can be used for social life and for overcoming problems encountered in daily life.

Mastery of reading literacy has become one of the essential skills to be acquired in the 21st century. Life in this century is faced with challenges in the form of technological development that no longer follows a linear pattern but rather an exponential pattern [2]–[5]. Various technologies that are developing in this day and age are artificial intelligence (AI), the internet of things (IoT), cloud computing and big data. In addition, the technologies are supported by the cyber-physical system (CPS), a system that allows a set of devices to interact and integrate with the real world [6], [7]. The rapid development of technology is thought to have

affected the loss and emergence of various types of jobs. Therefore, mastering reading literacy is important so that humans are able to compete with high-tech machines and are able to utilize what they read to anticipate future changes.

It is unfortunate that the urgency of the literacy is not supported by the literacy achievements of students in Indonesia. The results of the Program for International Student Assessment (PISA) measurement indicated that students' literacy achievement in Indonesia was at the 72nd level out of 77 countries. The average score achieved was 371, while the average reading literacy achievement of Organisation for Economic Co-operation and Development (OECD) member students was 487 [8]. In addition, the measurement of Progress in International Reading Literacy Skills (PIRLS) shows that Indonesian students' literacy achievement was ranked 42nd out of 45 countries in 2011 [9] and 41st out of 45 countries in 2006 [10]. Based on these two international measurements, it is known that the literacy achievement of Indonesian students is still low and almost at the bottom of the list of countries participating in these two types of measurements.

With regard to international measurements and based on research conducted in Chile, Costa Rica and Mexico, it was known that PISA 2015 contained items that were biased toward test takers from these three countries [11]. Furthermore, the S12 test booklet of PISA 2015 also contains bias against test takers from Australia, France, Singapore, and Turkey [12]. The weaknesses of this international test could be caused by translation, the cultural group of test takers, the familiarity of test takers cultural group with the content of the test items, the use of expressions and the test format [12], [13]. A set of tests that is not free of bias can lead to measurement results that are unable to truly explain the literacy achievements of the test takers. Thus, it is necessary to develop a test that is free of bias, so that it is able to produce measurements that reflect the actual literacy achievements of the test takers, including in the context of Indonesia. In this regard, in 2018, Indonesian government developed an instrument to measure reading literacy achievement. This instrument was developed by linguists and measurement experts from two prestigious institutions under the Ministry of Education, Culture, Research, and Technology of the Republic of Indonesia, namely the Center for Language Strategy and Diplomacy Development as well as the Educational Assessment Center. The instrument includes Indonesian content. This is expected to minimize the challenges that arise from international tests.

However, it is necessary to verify the quality of the instrument. This needs to be done to determine its quality. It is expected that the instrument is of good quality so that it only causes low a measurement error. In other words, it is expected that measurements using the instrument will produce scores that can reflect the actual abilities of the test takers. A test is considered to be good when its validity is verified and it has a high reliability coefficient, among others. Validity refers to the functioning of the instrument to measure the abilities or skills to be measured [14], [15]. Thus, verifying the validity is important to determine whether or not the instrument can be used to measure students' reading literacy skills. In addition, estimating the reliability is no less important. Reliability is the consistency of the value of the measurement results [16], [17]. Therefore, the higher reliability coefficient of an instrument, the better it reflects the consistency of the measurement results.

As previously explained, validity and reliability are important to determine the accuracy and constancy of the developed test in measuring the aspects to be measured. Verifying the validity and estimating the reliability of the test are important to provide an empirical basis for observers, researchers and literacy-related policymakers who will use the test. In addition, for researchers in the field of educational measurement and evaluation, the results of this study can enrich or inspire them in developing literacy-related tests, especially reading literacy.

2.   METHOD

Verification of the validity and reliability of the test for measuring students' reading literacy achievement was carried out based on the responses of 6,539 students from 294 high schools in 34 provinces in Indonesia. The sample was established using a multistage sampling technique from the study population of 15-year-old students in Indonesia, who had almost completed compulsory education and acquired the knowledge and skills to be able to participate in society [18]. The total number of respondents was 4,034 female respondents and 2,505 male respondents. The mean age of the respondents was 15 years with a standard deviation of 1.3. Based on school status, 3,864 respondents were students of public schools and 2,675 respondents were students of private schools. Based on the results of the national examination, 2,144 respondents were students of schools with national examination results belonged in the good category, 2,212 respondents were students of schools with national examination results in the moderate category and 2,183 respondents were students of schools with national examination results in the poor category.

The set of tests under study consisted of 40 questions which were developed by the Center for Language Strategy and Diplomacy Development and the Educational Assessment Center. The institutions are under the auspices of the Ministry of Education, Culture, Research, and Technology of the Republic of Indonesia. In it is development, the test items were packaged according to PISA framework. Therefore, the development of a set of tests considered the dimensions of reading literacy in PISA, which included the format of reading materials, types of reading tasks, and reading situations or contexts [19]. As mentioned earlier, this research aims to verify the quality of a set of tests using secondary data, not to discuss its developmental stages. Details of text titles, number of questions and codes for the nine texts that make up the set of tests are shown in the following Table 1.

Table 1. Text details

No	Text titles	Total number of items	Code
1	*Batik: Budaya Indonesia yang mendunia* (Batik: Indonesian culture that goes global)	7	*Batik* (Batik)
2	*Laskar pelangi* (Rainbow warriors)	6	*Laskar* (Warriors)
3	*Surat izin tidak masuk sekolah* (School absence permit)	3	*Penyakit Vektor*
4	*Lima provinsi tertinggi angka insiden DBD* (Five provinces with the highest dengue hemorrhagic fever (DHF) incidence rate)	2	(Disease)
5	*Fitur pertanyaan sering muncul (PSM) di Laman sehatitupenting.com* (FAQ feature on page sehatitupenting.com)	4	
6	*Perbandingan musim* (Season comparison)	6	*Musim* (Season)
7	Amanda *dan sang ratu* (Amanda and the queen)	4	*BIN* (BIN)
8	*Pengunjung perpustakaan* (Library visitors)	4	
9	*Rencana Internasional* (International plans)	4	

This measurement does not focus on lexical or syntax knowledge but rather on students' ability to engage with and utilize written information. More specifically, students are assessed on their ability to accurately use written information with particular topics for specific purposes. It is in line with the urgency of reading literacy mastery in this century, which is to utilize information for specific purposes.

Furthermore, the verification of validity focused on the topics covered in each text within this set of tests. This is important to do in order to show that the test items truly measure students' reading literacy achievements according to the topics of each text. The construct validity was verified using confirmatory factor analysis (CFA). CFA is a type of factor analysis that models relationships between latent variables and the observed variables [20], [21]. In other words, CFA is often used to test the construct validity of a set of tests by checking whether the latent variables underlying a set of indicators align with the theoretical conceptualization of the constructs that are intended to measure [21], [22]. In this study, CFA was chosen to confirm the construct of the observed variables in measuring the latent variable, which was the reading literacy achievement of Indonesian students. Based on the constituent text, the construct of this measurement instrument is shown in Figure 1.
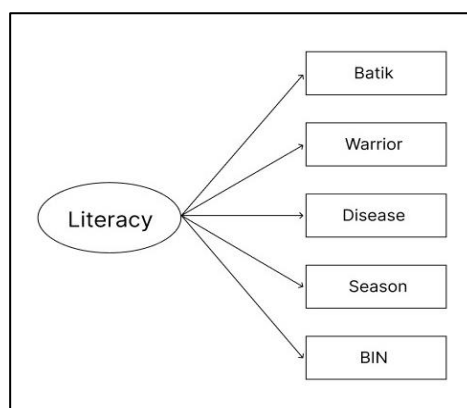

Figure 1. The construct of the students' reading literacy attainment test

The verification was preceded by proving suitability of the model with the student's response. Suitability of the model was determined by comparing the calculated values with the criteria. The various calculated values included the p-value chi-square, root mean square error of approximation (RMSEA), comparative fit index (CFI), tucker-lewis index (TLI) and standardized root mean square residual (SRMR) with criteria, among others.

After the compatibility of the model was fulfilled, the researcher examined thoroughly the value of the factor loading to determine the extent to which the observed variables could represent latent variables. The factor loading value is expected to reach a value above 0.3 because this value is the minimum limit for manifest factor loading or indicators that can represent latent [23]. As known, a factor loading value of 0.3 indicates a moderate relationship between two variables [24], [25]. Furthermore, the value of the factor loading was used to estimate the reliability of the items. A test has a good reliability value if the calculation results produce a minimum value of 0.7 [26]–[28]. The estimation was carried out using the ω as in (1) [29], [30].

$$\omega = \frac{\left(\sum_{i=1}^{k} \lambda_i\right)^2}{\left(\sum_{i=1}^{k} \lambda_i\right)^2 + \left(\sum_{i=1}^{k} \delta i\right)} \tag{1}$$

3. RESULTS AND DISCUSSION
3.1. Test for measuring students' reading literacy achievement

Before discussing the validity and reliability of the test for measuring Indonesian students' reading literacy achievement in more detail, this section will first discuss the texts that make up this measurement instrument. However, details of the texts and accompanying questions can be found in *Laporan kajian bahan kebijakan teknis literasi nasional* (National literacy technical policy material study report) published by the Agency for Language Development and Cultivation in 2018 [19].

As previously mentioned, there are nine texts that make up the test for measuring Indonesian students' reading literacy achievement. The first text is coded Batik and entitled *Batik: Budaya Indonesia yang mendunia* (Batik: Indonesian culture that goes global). As the title suggests, the text discusses one of Indonesia's cultural assets, namely *Batik*. The discussion in the text focuses on the popularity of *Batik* at the international level by presenting supporting arguments for this popularity. The type of the text is exposition, which states a point of view by presenting supporting evidence [31]. Exposition texts have the potential to enhance students' understanding and to teach them how to express opinions in providing information on a phenomenon [32]. There are 10 paragraphs that make up this first text. The constituent paragraphs are in accordance with the structure of the expository text, namely introduction, series of arguments and closing [33].

The following text is an excerpt from a novel written by a well-known writer in Indonesia, namely Andrea Hirata. The novel is entitled *Laskar pelangi* (Rainbow warriors). The code for this text and the accompanying question items is *Laskar* (Warriors). *Laskar pelangi* (Rainbow warriors) belongs to the type of text commonly read by general public, i.e., narrative text [31]. A narrative is a form of written work, often fictional, in which a character confronts a challenge and attempts to solve it by navigating a sequence of occurrences [34]. Apart from novels, there are other various forms of writing categorized as narrative text; among others, short stories, fairy tales and autobiographies [35], [36]. In daily life, narrative texts can be found everywhere and make the world more meaningful through storytelling [37]. Regarding it is usefulness, the narrative text plays a big role in social life. The text has the potential to become a medium that can be used to change social opinions [31]. In terms of text structure, there are three components that make up the narrative text, namely orientation, sequence of events and resolution [31]. In the excerpt from the novel *Laskar pelangi* (Rainbow warriors), through the orientation section in the first paragraph, the reader can obtain information regarding the background of events and characters. Broadly speaking, it is told in the paragraph that one of the characters is preparing a school disbandment speech. The speech is prepared due to his concern that the school will not be able to get enough students and thus, the school will have to be disbanded. Furthermore, in the sequence of events, the events that occurred is told according to the time sequence of the story. At the end of the story, there is a resolution paragraph describing the feelings of the characters. Through this depiction, the readers know that the story has a happy ending. The school remains open because there are even ten new students who will study at the school.

In addition to texts on educational themes, the instrument contains texts that discuss health. Code for the text is *penyakit vektor* (vector-borne disease/disease). There are three texts with this code, namely: i) *Surat Izin tidak masuk sekolah* (School absence permit), ii) *Lima Provinsi tertinggi angka insiden DBD* (Five provinces with the highest DHF incidence rate), and iii) *Fitur pertanyaan sering muncul (PSM) di laman sehatitupenting.com* (Frequently asked questions feature on page sehatitupenting.com). The first text contains a

permission letter written by parents addressed to a homeroom teacher. Through the letter, the parents ask the teacher to allow their child not to attend classes due to dengue fever. In addition, a text describing the permission letter is presented even before the letter. It is a descriptive text, a type of text that functions to describe something specifically [33]. Through the text, test takers obtain information about permits, including the meaning, benefits and media for writing a permission letter. The next text is entitled *Lima provinsi tertinggi angka insiden DBD* (Five provinces with the highest DHF incidence rate). The text contains information about five provinces with the highest number of DHF incidents. First, data on the five provinces with the highest DHF rates in 2005-2009 are presented in graphical form accompanied by writing that supports the graphic presentation. Second, data on provinces with the highest DHF incidence rates in the 2009-2013 range are presented in tabular form. The last text with the *Penyakit vektor* (Vector-borne disease) code contains an explanation of the dengue fever cycle. An explanation of the dengue fever cycle is not only explained in written text but also presented in graphics that help readers understand. The text is of the explanatory type that explains a series of events or phenomena [33].

Next is the text with the title *Perbandingan musim* (Season comparison) with *perbandingan* (comparison) as the code. The text compares the seasons in two countries in the Southeast Asia Region, namely Thailand and Indonesia. The comparison is presented in two texts; the first is about the seasons in Thailand and the second is about the seasons in Indonesia. In line with the text coded Batik, *perbandingan musim* (peason comparison) text is an exposition text. Aside from paragraphs explaining the seasons in the two countries previously mentioned, the text is accompanied by pictures that show the country's identity and a description of the seasons. Through the text, test takers get information about the seasons in Thailand and Indonesia. Other various important information that can be obtained from the text includes the best months to visit the two countries and transitional periods that cause many diseases, such as flu and coughs.

Finally, there are three different texts under code BIN: i) Amanda *dan sang ratu* (Amanda and the queen), ii) *Pengunjung perpustakaan* (Library visitors), and iii) *Rencana internasional* (International plans). The first text, Amanda *dan sang ratu* (Amanda and the queen), belongs to drama genre which is one of the literary works, apart from poetry and prose [38], [39]. Drama scripts contain dialogues, not narration [40]. Judging from it is content, the text of the drama script contains spoken language which is arranged in the form of dialogue acted out by the actors in front of drama lovers. In the test, apart from the dialogues, Amanda *dan sang ratu* (Amanda and the queen) includes definitions of various theatrical jobs, such as actor, director and sound engineer. Thus, through the text, test takers do not only enjoy literary work but also obtain additional information related to several theater works. The second text, i.e., *Pengunjung perpustakaan* (Library visitors), is in the form of an image accompanied by text. In the text, two maps of the library are presented accompanied by writing about items located in the library, such as bookshelves and their types of books, information desks, circulation desks and access in and out of the library. In line with the text about the library, the text entitled *Rencana internasional* (International plans) is a continuous text, a text which is a combination of pictures and writing [18]. The third text is in the form of a graphic supplemented with writings related to graphics. The graphs present information on various international programs being carried out in East and South Africa. On the whole, there are three main programs implemented, namely programs on healthy living, learning and habitat.

3.2. Validity and reliability of the test for measuring reading literacy achievement

Verification of the validity of the instrument for measuring Indonesian students' reading literacy achievement was carried out globally for the entire instrument using CFA. CFA can be conducted using different approaches, including item-level approach, second-order approach and testlet-based or item parceling approach [41]. Testlet-based or item parceling was used in this research. First, test items were combined to form a set of values which were then analyzed [42], [43]. Verification of the validity was preceded by determining the suitability of the model and followed by examining values of factor loading. The calculated values obtained for proving model suitability are presented in Table 2.

Table 2. Calculation results, criteria and conclusions of the fit of the model

Goodness of fit	Criteria	Score	Category
P-Value	>0.05 [44]	0.056	Fit
RMSEA	≤0.06 [45]	0.034	Fit
CFI	Fit: CFI≥0.95 [45] or Accepted: CFI 0.90-0.94 (Bentler in [45])	0.995	Fit
TLI	Fit: TLI≥0.95 [45] or Accepted: TLI 0.90-0.94 (Bentler in [45])	0.991	Fit
SRMR	<0.08 [46]	0.017	Fit

Table 2 shows that the construct model of the instrument is supported by student response patterns as empirical data. This is because the calculated value meets the criterion value. Furthermore, through the results of data analysis, the value of factor loading was obtained. The Figure 2 displays construct profile of the instrument for measuring students' reading literacy achievement based on the results of CFA towards student response patterns on the test.

Figure 2 shows the factor loadings connecting the latent variable (reading literacy) with the observed variable Batik, warriors, disease, season, and BIN respectively in their order are 0.557; 0.767; 0.583; 0.548; and 0.761. All of the factor loadings have a value above 0.3, which is the minimum limit for manifest factor loading or indicators that can represent latent [22], [23]. Thus, based on the CFA described previously, it is known that the instrument for measuring Indonesian students' reading literacy achievement has proven to have construct validity.
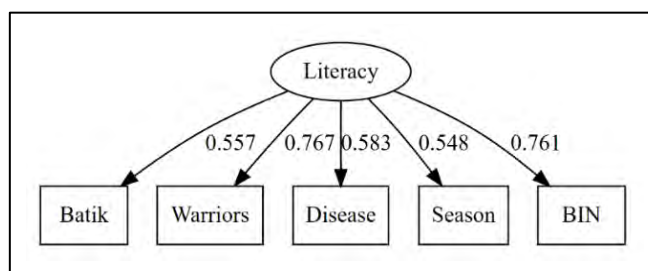

Figure 2. The test construct for measuring students' reading literacy achievement

Then, the reliability value of the instrument was estimated using the factor loading value depicted in Figure 2. The estimation was carried out using equation (1). The calculation results show that the test has a reliability value of 0.805. Thus, it can be seen that the reliability value is greater than the minimum value for good reliability, which is 0.70 [27], [28]. More specifically, the acceptable reliability value based on calculations using the omega formula is 0.65-0.8 [47]. In other words, the test has good stability in measuring students' reading literacy achievement.

Judging from the results of verifying the validity and estimating the reliability, the instrument for measuring Indonesian students' reading literacy achievement has good characteristics. This means that the test is appropriate to be used to measure Indonesian students' reading literacy achievement. Indeed, the instrument is of high quality and can accurately measure reading literacy achievement because the set of tests was developed carefully by linguists and measurement experts by taking into account various factors in creating a good test, among others, by paying attention to text selection criteria. Various criteria must be considered when selecting texts for the preparation of tests, namely variety, suitability of the content, exploitability, text length, readability, originality and presentation [48]-[51]. On this occasion, we will discuss the texts that make up the students' reading literacy achievement test based on the two criteria for selecting a good text, namely variety and appropriateness of the content.

The variety of texts can be seen from the various types of texts used [49]. Based on the previous description, it is known that the nine texts that make up the instrument for students' reading literacy achievement have various text types, including exposition, narrative, description and explanation. In addition, the instrument contains texts in the form of drama scripts, permits, reports, research results in graphic and chart formats, as well as room plans [19]. These various text types have been studied by Indonesian students, even at the elementary education level. At the lowest level of formal education, namely elementary school, students have studied various types of texts, such as descriptive texts, narrative texts, informative report texts, directive/instructive texts and explanatory texts [52]. Furthermore, at the junior high school level, students restudy various types of text as shown in Table 3 [53]. Table 3 shows that junior high school students learn various types of texts that are tested in measuring their reading literacy achievement. In addition, students have also studied other various forms of texts that are also tested, namely letters and plays.

Then, the suitability of the content can be reviewed on the topic of the text. The topics raised by the text are appropriate for students in Indonesia. There is no doubt that the text entitled *Batik: Budaya Indonesia yang mendunia* (Batik: Indonesian culture that goes global) raises a topic that is very suitable for Indonesian students. *Batik* plays a significant role in Indonesian culture, seen as more than just a product of Indonesian society's creativity [54]. The United Nations Educational, Scientific and Cultural Organization (UNESCO) has given recognition to *Batik* as an intangible cultural heritage of humanity belonging to Indonesia. In order

to appreciate *Batik* as a cultural heritage, the government of the Republic of Indonesia established October 2nd as Batik Day [55]. Thus, *Batik* has become an inseparable part of the daily lives of Indonesian people, including students from all over the archipelago. In addition, the text entitled *Laskar pelangi* (Rainbow warriors) is no less popular among Indonesians, including students. Through this novel, readers can get an overview of the lives of Indonesian people, especially the lives of people in Belitong, Bangka Belitung Islands. Furthermore, the text with the code *Penyakit vektor* (Vector-borne disease/disease) contains a topic that is familiar to students, namely dengue fever. The term vector-borne disease is relatively rare in everyday life. However, several diseases included in this group have become popular in the lives of Indonesian people, including students in Indonesia. Vector-borne disease is a disease that is transmitted through vectors and disease-carrying animals. Some examples of vector-borne diseases are Zika fever, Malaria, dengue fever, West Nile fever, and Chikungunya [56]–[58]. The disease is transmitted by vectors and disease-carrying animals, including mosquitoes, flies, cockroaches and rats [59]. More specifically, the text coded as *Penyakit vektor* (Vector-borne disease/disease) raises a topic of dengue fever. In Indonesia, the annual rate of DHF cases increased sharply over the 50-year period, from 1968 to 2016 [60]. One of the places that has the potential for the spread and transmission of DHF is school. Therefore, the government and school institutions are making eradication efforts by carrying out various activities, such as *menguras*/draining, and *mengubur*/burying the secondhand, *menutup*/closing the water storage, and *memantau*/monitoring the breeding site (4M) [61].

Table 3. Scope of Indonesian language materials for junior high schools

Year VII	Year VIII	Year IX
Description	News	Trial Report
Fantasy Story	Advertisement	Speech
Procedure	Exposition	Short story
Observation Report	Poetry	Response
People's Poetry	Explanation	Discussion
Folklore	Review	Inspirational Stories
Letter	Persuasion	Literacy
Literacy	Drama	
	Literacy	

4. CONCLUSION

Based on the results of data analysis, it is known that the instrument for measuring student' reading literacy achievement in Indonesia is proven to be valid and has a good reliability value. In other words, the instrument is appropriate for measuring students' reading literacy achievement in Indonesia. As previously mentioned, the construct validity in this study focuses on the topic of the text. Subsequent research should be focused on the construct validity based on the process dimensions of reading literacy, namely accessing and retrieving, integrating and interpreting as well as reflecting and evaluating. Besides that, as presented in the findings and discussions section, the variety and suitability of text content, two of the seven criteria that must be considered in selecting a text for a set of tests, have been explored in the study. Furthermore, deep exploration can be carried out regarding the characteristics of each text composing the test in terms of exploitability, text length, readability, originality and presentation. In addition, researchers in the field of measurement can explore more deeply the characteristics of test items based on modern test theory. This is done to determine the difficulty level and the discrimination power of each item. Estimating the characteristics of the test items is a step in collecting item banks related to reading literacy. Furthermore, in order to enrich the reading literacy item bank, future research can focus on developing items to measure students' reading literacy achievement for other levels of education, such as elementary and junior high school levels. This will be very useful for measuring the literacy achievement of Indonesian students from an early age.

REFERENCES

[1]   S. B. Kucer, *Dimensions of literacy: a conceptual base for teaching reading and writing in school settings,* fourth edition. Routledge, 2014.

[2]   J. Hagel, J. S. Brow, T. Samoylova and M. Lui, "From exponential technologies to exponential innovation," Deloitte University Press, 2013.

[3]   M. Yamin, "Information technologies of 21st century and their impact," *International Journal of Information Technology*, vol. 11, pp. 759-766, 2019, doi: 10.1007/s41870-019-00355-1.

[4]   R. S. Malik, "Educational challenges in 21st century and sustainable development," *Journal of Sustainable Development Education and Research*, vol. 2, no. 1, pp. 9-20, 2018, doi: 10.17509/jsder.v2i1.12266.

[5]   A. Bohr and K. Memarzadeh, "The rise of artificial intelligence in healthcare applications," in *Artificial Intelligence in Healthcare*, Elsevier, 2020, pp. 25–60, doi:10.1016/B978-0-12-818438-7.00002-2.

[6]   N. Boulila, "Cyber-physical systems and Industry 4.0: properties, structure, communication, and behavior," *Siemens Corporate Technology, Munich*, 2019, doi: 10.13140/RG.2.2.27890.76485.

[7]   R. Wisniewski, G. Bazydło, P. Szcześniak, I. Grobelna, and M. Wojnakowski, "Design and verification of cyber-physical systems specified by petri nets-a case study of a direct matrix converter," *Mathematics*, vol. 7, no. 9, pp. 812, Sep. 2019, doi: 10.3390/math7090812.

[8]   A. Schleicher, "PISA 2018: Insights and interpretations," *OECD Publishing*, 2019.

[9]   S. Thompson *et al.*, "Highlights from PIRLS 2011 reading achievement of U.S. fourth-grade students in an International Context," *National Center For Education Statistics*, pp. 1–64, 2012.

[10]   J. Baer, S. Baldi, K. Ayotte, and P. J. Green, "The reading literacy of U.S. fourth-grade students in an international context results from the 2001 and 2006 progress in international reading literacy study (PIRLS)," U. S. Department of Education NCES 2008-017, pp. 1–69, 2007.

[11]   C. Dzul-Garcia and B. Atar, "Investigation of possible item bias on PISA 2015 science items across Chile, Costa Rica and Mexico," *Culture and Education*, vol. 32, no. 3, pp. 470–505, Jul. 2020, doi: 10.1080/11356405.2020.1785158.

[12]   B. Alatlı, "Cross-cultural measurement invariance of the items in the science literacy test in the programme for international student assessment (PISA-2015)," *International Journal of Education and Literacy Studies*, vol. 8, no. 2, May 2020, doi: 10.7575/aiac.ijels.v.8n.2p.16.

[13]   B. Alatlı, "An investigation of cross-cultural measurement invariance and item bias of PISA 2018 reading skills items," *Journal of Education and Teaching (IOJET)*, pp. 1047–1073, 2022.

[14]   G. D. Garson, "Validity and reliability." *Asheboro, NC: Statistical Associates Publishers*, pp. 9–28, 2013.

[15]   S. J. Kamper, "Reliability and validity: linking evidence to practice," *Journal of Orthopaedic and Sports Physical Therapy*, vol. 49, no. 4, pp. 286–287, Apr. 2019, doi: 10.2519/jospt.2019.0702.

[16]   L. M. Crocker and J. Algina, *Introduction to classical and modern test theory*. ERIC, 1986.

[17]   L. Sürücü and A. Maslakçi, "Validity and reliability in quantitative research," *Business and Management Studies: An International Journal*, vol. 8, no. 3, pp. 2694–2726, Sep. 2020, doi: 10.15295/bmij.v8i3.1540.

[18]   OECD, *PISA 2018 Assessment and analytical framework,* Paris: OECD Publishing, 2018, doi: 10.1787/b25efab8-en.

[19]   *Badan Pengembangan dan Pembinaan Bahasa* (The Agency for Language Development and Cultivation), "*Laporan kajian bahan kebijakan teknis literasi nasional* (Report on the study of national literacy technical policy materials)." Jakarta: Badan Pengembangan dan Pembinaan Bahasa, 2018. (Accessed date: April 19, 2024), [Online], Available at https://billyantoro.com/laporan-kajian-bahan-kebijakan-teknis-literasi-nasional-tahun-2018/

[20]   J. M. Roos and S. Bauldry, *Confirmatory factor analysis*. Sage Publications, 2021.

[21]   C. D. Nye, "Reviewer resources: confirmatory factor analysis," *Organizational Research Methods*, vol. 26, no. 4, pp. 608–628, Oct. 2023, doi: 10.1177/10944281221120541.

[22]   R. Gomez, V. Stavropoulos, and M. D. Griffiths, "Confirmatory factor analysis and exploratory structural equation modelling of the factor structure of the depression anxiety and stress Scales-21," *Plos One*, vol. 15, no. 6, Jun. 2020, doi: 10.1371/journal.pone.0233998.

[23]   J. F. Hair, R. E. Anderson, B. J. Babin, and W. C. Black, *Multivariate data analysis: A global perspective*. Upper Saddle River, NJ: Pearson, 2010.

[24]   M. Tavakol and A. Wetzel, "Factor analysis: a means for theory and instrument development in support of construct validity," *International Journal of Medical Education*, vol. 11, pp. 245–247, Nov. 2020, doi: 10.5116/ijme.5f96.0f4a.

[25]   J. Li, T. Sawaragi, and Y. Horiguchi, "Introduce structural equation modelling to machine learning problems for building an explainable and persuasive model," *SICE Journal of Control, Measurement, and System Integration*, vol. 14, no. 2, pp. 67–79, Jun. 2021, doi: 10.1080/18824889.2021.1894040.

[26]   G. W. Cheung, H. D. Cooper-Thomas, R. S. Lau, and L. C. Wang, "Reporting reliability, convergent and discriminant validity with structural equation modeling: A review and best-practice recommendations," *Asia Pacific Journal of Management*, Jan. 2023, doi: 10.1007/s10490-023-09871-y.

[27]   G. Groth-Marnat and A. J. Wright, *Handbook of psychological assessments*, 6th ed. New Jersey: Wiley, 2016.

[28]   N. K. Chadha, *Applied psychometry*. SAGE Publications Pvt. Limited, 2009.

[29]   G. J. Geldhof, K. J. Preacher, and M. J. Zyphur, "Reliability estimation in a multilevel confirmatory factor analysis framework," *Psychological Methods*, vol. 19, no. 1, pp. 72–91, 2014, doi: 10.1037/a0032138.

[30]   M. A. Padilla and J. Divers, "A comparison of composite reliability estimators," *Educational and Psychological Measurement*, vol. 76, no. 3, pp. 436–453, Jun. 2016, doi: 10.1177/0013164415593776.

[31]   P. Knapp and M. Watkins, *Genre, text, grammar*. unsw Press, 2005.

[32]   A. Kemala, A. Sastromiharjo, and Y. Mulyati, "Students' problems in writing exposition texts: structure, function, and effectiveness," 2020, doi: 10.2991/assehr.k.200406.003.

[33]   D. Rose and J. R. Martin, *Learning to write, reading to learn: Genre, knowledge and pedagogy in the Sydney School*. Equinox London, 2012.

[34]   E. Hill, A. Whitworth, M. Boyes, M. Ziegelaar, and M. Claessen, "The influence of genre on adolescent discourse skills: Do narratives tell the whole story?," *International Journal of Speech-Language Pathology*, vol. 23, no. 5, pp. 475–485, Sep. 2021, doi: 10.1080/17549507.2020.1864016.

[35]   N. Yazıcı, İ. A. Gürsoy and G. D. D. Aydemir, "Aspects of narrative competence in language levels and narrative texts in CEFR," *Kastamonu Education Journal*, vol. 29, no. 3, pp. 559-573, 202, doi: 10.24106/kefdergi.774173.

[36] R. A. Mar, J. Li, A. T. P. Nguyen and C. P. Ta, "Memory and comprehension of narrative versus expository texts: a meta-analysis," *Psychonomic Bulletin & Review*, vol. 28, pp. 732-749, 2021, doi: 10.3758%2Fs13423-020-01853-.

[37] B. Forchtner, "Critique, Habermas and narrative (genre): the discourse-historical approach in critical discourse studies," *Critical Discourse Studies*, vol. 18, no. 3, pp. 314–331, May 2021, doi: 10.1080/17405904.2020.1803093.

[38] M. I. Amrullah, T. N. I. Wahya and W. I. M. Rohmana, "Drama as the way to integrate literature in English language teaching," *International Journal of English Learning and Applied Linguistic*, vol. 3, no. 2, pp. 88-94, 2023, doi: 10.21111/ijelal.v3i2.9057.

[39] G. Chen, "A literature review on prose study," *Advances in Social Science, Education and Humanities Research*, vol. 378, pp. 512-516, 2019, doi: 10.2991/assehr.k.191217.117.

[40] A. Beard, *Text and contests*. New York: Routledge, 2001.

[41] A. D. Hapsari and W. Widhiarso, "The effectiveness of item parceling to increase the model fit: A case study of PAPs," *Jurnal Penelitian dan Evaluasi Pendidikan*, vol. 27, no. 1, pp. 26–38, Jun. 2023, doi: 10.21831/pep.v27i1.49012.

[42] B. Thompson and J. G. Melancon, *Using item 'testlest'/'parcels' in confirmatory factor analysis: An example using the PPSDQ-78*. 1996.

[43] T. D. Little, W. A. Cunningham, G. Shahar, and K. F. Widaman, "To parcel or not to parcel: exploring the question, weighing the merits," *Structural Equation Modeling: A Multidisciplinary Journal*, vol. 9, no. 2, pp. 151–173, Apr. 2002, doi: 10.1207/S15328007SEM0902_1.

[44] R. B. Kline, *Principles and practice of structural equation modelling*, 2 ed., New York: The Guilford Press, 2005.

[45] T. A. Brown, *Confirmatory factor analysis for applied research*, vol. 44, no. 05. Guilford Publications, 2015.

[46] K. Gana and G. Broc, *Structural equation modeling with lavaan*. New Jersey: Wiley, 2019.

[47] L. Nikolayeva, "Adapting a text for testing purposes: approach to academic reading and writing assessment design in the Arab world," *Theory and Practice in Language Studies*, vol. 9, no. 2, Feb. 2019, doi: 10.17507/tpls.0902.10.

[48] M. T. Kalkbrenner, "Alpha, omega, and H internal consistency reliability estimates: reviewing these options and when to use them," *Counseling Outcome Research and Evaluation*, vol. 14, no. 1, pp. 77–88, Jan. 2023, doi: 10.1080/21501378.2021.1940118.

[49] Suyadi and Husnaini, "Reading skills used in English Textbook," *Jurnal Ilmiah Universitas Batanghari Jambi*, vol. 22, no. 2, pp. 1322-1328, 2022, doi: 10.33087/jiubj.v22i2.2405.

[50] S. A. Berardo, "The use of authentic materials in the teaching of reading," *The Reading Matrix: An International Online Journal*, vol. 6, no. 2, pp. 60–69, 2006.

[51] S. Bibby, "Criteria and creation: Literary texts and a literature textbook," *The Journal of Literature in Language Teaching*, vol. 3, no. 1, pp. 20–31, 2014.

[52] "*Kementerian Pendidikan dan Kebudayaan* (Ministry of Education and Culture), Kurikulum 2013: *Kompetensi Dasar Sekolah Dasar/Madrasah Ibtidaiyah* (Kurikulum 2013: Basic competencies for elementary schools/Islamic elementary schools)." Jakarta: Kementerian Pendidikan dan Kebudayaan, 2013, (Accessed date: April 19, 2024). [Online], Available at https://www.academia.edu/4951613/KOMPETENSI_DASAR_Sekolah_Dasar_SD_Madrasah_Ibtidaiyah_MI_

[53] "*Kementerian Pendidikan dan Kebudayaan* (Ministry of Education and Culture), *Materi bimbingan teknis fasilitator dan instruktur Kurikulum 2013: SMP mata Pelajaran Bahasa Indonesi*a (Technical guidance material for Curriculum 2013 facilitators and instructors." Jakarta: Kementerian Pendidikan dan Kebudayaan, 2017. (Accessed date: April 19, 2024). [Online], Available at https://lpmpdki.kemdikbud.go.id/wp-content/uploads/2020/02/K13-B.Indonesia.pdf.

[54] [S. Bin Silah, B. Isa, Ponimin and T. Karyono, "Malaysian and indonesian batik: the best practices," *International Conference on Art, Design, Education and Cultural Studies (ICADECS)*, pp. 47-55, 2021, doi: 10.18502/kss.v5i6.9177.

[55] S. Sumardi, A. Ayu, S. Kantun and M. Mashudi, "Batik social and cultural legitimacy globalization: study analysis and recognition Batik Motif Jember through Jember Fashion Carnival," *Advance in Social Science, Education and Humanities Research*, vol. 125, pp. 53-57, 2018, doi: 10.2991/ICIGR-17.2018.13.

[56] B. Chala and F. Hamde, "Emerging and re-emerging vector-borne infectious diseases and the challenges for control: a review," *Frontiers in Public Health*, vol. 9, Oct. 2021, doi: 10.3389/fpubh.2021.715759.

[57] M. C. Thomson and L. R. Stanberry, "Climate change and vectorborne diseases," *New England Journal of Medicine*, vol. 387, no. 21, pp. 1969–1978, Nov. 2022, doi: 10.1056/NEJMra2200092.

[58] J. Rocklöv and R. Dubrow, "Climate change: an enduring challenge for vector-borne disease prevention and control," *Nature Immunology*, vol. 21, no. 5, pp. 479–483, May 2020, doi: 10.1038/s41590-020-0648-y.

[59] T. S. Athni *et al.*, "The influence of vector-borne disease on human history: socio-ecological mechanisms," *Ecology Letters*, vol. 24, no. 4, pp. 829–846, Apr. 2021, doi: 10.1111/ele.13675.

[60] H. Harapan, A. Michie, M. Mudatsir, R. T. Sasmono and A. Imrie, "Epidemiology of dengue hemorrhagic fever in Indonesia: analysis of five decades data from the National Disease Surveillance," *BMC Research Notes*, vol. 12, no. 1, pp. 1-6, 2019, doi: https://doi.org/10.1186/s13104-019-4379-9.

[61] V. L. Amelia, A. Setiawan and Sukihananto, "Board game as an educational media for dengue prevention knowledge for school-aged children," *Enfermeria Global*, vol. 18, no. 56, pp. 264-272, 2019, doi: 10.6018/eglobal.18.4.336611.

BIOGRAPHIES OF AUTHORS

Limala Ratni Sri Kharismawati is currently doing her doctoral program in Educational Research and Evaluation, Yogyakarta State University. Besides, she also works as Head of Division of Research and Development Program, Southeast Asian Ministers of Education Organization (SEAMEO) Regional Centre for Quality Improvement of Teachers and Education Personnel (QITEP) in Language/SEAQIL. Her interest includes teaching German as a foreign language, literacy and educational assessment. She can be contacted at email: limala.ratni2016@student.uny.ac.id or limala.ratnisk@yahoo.de.

Pratomo Widodo (iD) 🔳 SC ⬡ is a professor in Department of German Language Education, Faculty of Languages, Arts and Cultures, Yogyakarta State University. He also teaches at the Graduate School, Yogyakarta State University. His research interests include applied linguistics, German studies, and teaching German as foreign language. Currently, he is Head of Asosiasi Germanistik Indonesia (Indonesian German Studies Association). He can be contacted at email: pratomo@uny.ac.id or prat_wid@yahoo.com.

Heri Retnawati (iD) 🔳 SC ⬡ is a professor in Assessment in Department of Mathematics Education, Faculty of Mathematics and Natural Science, Yogyakarta State University. Her research interests include research on educational assessment and evaluation as well as literacy. She can be contacted at email: heri_retnawati@uny.ac.id.