

Construct Validation with Rasch Modeling of Pedagogical Literacy Measurement Instrument for Teachers in Junior High School Environment

 <https://doi.org/10.31004/jele.v10i1.639>

*Kholifatun Nisa Rahmi, Purwo Susongko, Suriswo ^{1,2,3} 

^{1,2,3} Universitas Pancasakti Tegas

Corresponding Author: kholifatunrahmi13@guru.smp.belajar.id

ABSTRACT

This study aims to develop and analyze a pedagogical literacy assessment instrument in the junior high school environment using the Rasch modeling. In this research, the development and analysis of the instrument involved an approach consisting of content, substantive, structural, and consequential aspects. The pedagogical literacy assessment instrument was developed based on ten teacher competency standards according to the Indonesian Ministry of National Education Regulation No. 16 of 2007, presented in the form of multiple-choice questions, totaling 60 items. This study is a Research and Development (R&D) with a research design using ADDIE (Analysis, Design, Development, Implementation, and Evaluation). The research approach used is mixed methods, which combines both qualitative and quantitative forms. Data was collected from 122 teachers across four schools: SMPN 7 Tegal City, SMPN 14 Tegal City, SMPN 17 Tegal City, and SMPN 19 Tegal City. The results of data analysis using the Rasch model on R programming version 4.03 with the eRm package showed that 88% of the items were valid with a p-value < 0.05. The validation results, which cover content, psychometrics, and constructs (content, substantive, structural, and external), met the requirements for being considered good items.

Keywords: Assessment, Pedagogical Literacy, Teachers In Junior High School Environment.

Article History:

Received 08th January 2025

Accepted 27th January 2025

Published 01st February 2025



INTRODUCTION

The mastery of a teacher's pedagogical competence can be seen from the ability of pedagogical literacy, which refers to the understanding and ability of an educator or teacher to apply effective educational concepts, theories and practices in the context of teaching. Pedagogical literacy is the ability to understand pedagogy, namely a thorough understanding of understanding teaching materials, knowing learners, and determining how to teach, teachers and the ability to integrate pedagogical knowledge into classroom learning practices. It is very important for teachers to provide effective and meaningful learning experiences for their students. For a teacher, pedagogical literacy helps teachers stay relevant in the dynamic world of education and improve the overall quality of education. To find out the pedagogical skills of teachers is currently academic supervision, academic supervision activities are equipped with supervision instruments only in the form of open-ended question lists and checklists, which describe the lesson planning that will be carried out by a teacher, there is no instrument in the form of a validated pedagogical literacy assessment model to determine the pedagogical literacy skills of teachers.

The concept of pedagogical literacy was developed from the standards of teachers' pedagogical competence in Pedagogical competence based on the Regulation of the Minister of National Education of the Republic of Indonesia Number 16 of 2007 concerning Standards

for Academic Qualifications and Teacher Competencies. Teachers' pedagogical competencies include 1. mastering student characteristics from physical, moral, social, cultural, emotional, and intellectual aspects; 2. mastering learning theories and principles of educational learning; 3. developing curriculum related to the field of development taught; 4. organizing educational development activities; 5. utilizing information and communication technology for the benefit of learning; 6. facilitating the development of students' potential to actualize various potentials. facilitating the development of students' potential to actualize their various potentials; 7. communicating effectively, empathetically, and politely with students; 8. conducting assessment and evaluation of learning processes and outcomes; 9. utilizing the results of assessment and evaluation for the benefit of learning; 10. taking reflective action to improve the quality of learning.

To produce a product to develop a pedagogical literacy assessment model for teachers in the junior high school environment, a development research is needed. A systematic process involving analysis, planning, development, experimentation and revision to produce a better and more effective product. Each development model has different characteristics. Some things that need to be considered to choose the type of development research model to be used are the purpose of the research, the type of research, and the resources available. A tool to find out from the assessment process using instruments, one of which is a written test. This written test aims to test teachers' understanding of pedagogical literacy. The test instrument must be carefully designed so that it can accurately measure what you want to measure. "Based on its significant role in research, the instrument must have certain qualifications that meet scientific requirements (Imania & Bariah, 2019) in Lestari et al., 2022.

So in this study, it is necessary to measure the validity so that the instrument actually measures according to the purpose of making the instrument, namely regarding the teacher's pedagogical competence. In the process of making a test, item analysis is needed with the aim of knowing the quality of the questions so that it can provide information about the characterization of the test items or empirical analysis (Sarea & Hadi, 2015) in (Mulyani et al., 2021). There are two types of item testing tests, namely classical theory and modern test theory, item review using Item Response Theory (IRT), one of which is used to analyze item tests. According to (Ayub et al., 2020) the Item Response Theory (IRT) approach, is a measurement approach developed to overcome the weaknesses of classical measurement theory. Classical measurement theory (classical true score model), among others, states that the characteristics of a test question depend on the group of test takers, the ability of the person observed as a performance value.

Based on (Indri Pudjiati & Faisal Madani, 2023) the Rasch model which is a type of Item Response Theory (IRT) determines the probability that a test taker can answer a test item correctly at a certain ability level. This model is based on item response theory which assumes that the ability of test takers can be measured accurately and objectively, and that the items used in the test have different levels of difficulty and produce scores that do not depend on the example of the question being worked on. The Rash model also calibrates three things: the measurement scale, test takers (people), and items. Calibration aims to ensure the validity and reliability of measurement results so that the test can provide comprehensive information (Ardiyanti, 2016) in (Pratama, 2020). Based on this opinion, the researcher used the Rasch model to analyze the items of the teacher pedagogical competency test to analyze validity.

The research objective can be formulated to prove the construct validity of the pedagogical literacy assessment model instrument for teachers in the junior high school environment.

METHOD

This research uses a type of development research or Research and Development (R&D). According to Sugiono in (Okpatrioka, 2023) research and development methods or research and development (R&D) in English refers to research methods used to produce certain products and test the effectiveness of these products. The development research model that researchers use to develop educational competency tests is the ADDIE model which is a mnemonic of Analysis Design Development Implementation and Evaluation. According to Sezer et al in (Rayanto & Sugianti, 2020). Several reasons the ADDIE model is still very relevant to use, namely (1) The ADDIE model is a model that is able to adapt well to various conditions and can still be used today; (2) Although the ADDIE model provides a high degree of flexibility in answering questions, the ADDIE model is an effective model and the ADDIE abbreviation is familiar to many people; (3) The ADDIE model provides a structured general framework for developing educational interventions and modifying and evaluating them at each stage Angko and Mustaji (2013: 4) in (Safitri & Aziz, 2022). The ADDIE model is schematized by Branch in (Hidayat & Nizar, 2021) in the following figure:

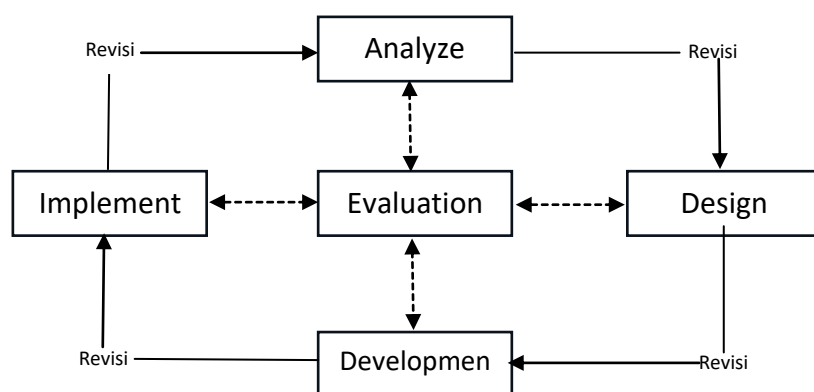


Figure 1. Stages of the ADDIE Model

The research subjects were teachers in the junior high school environment totaling 112 teachers in four junior high schools in Tegal City. The number of samples was calculated using the Slovin formula in (Agung Santoso, 2023) as follows

$$n = \frac{N}{1 + n(\mu)^2}$$

Description:

n : number of samples

N : total population

μ: margin of error, namely the amount of error that can be tolerated by researchers.

There are three data collection techniques carried out in this study, namely, first, data collection techniques using data collection questionnaires which are carried out by giving a set of written questions to respondents to answer which has a function so that researchers know exactly the variables to be measured and know what is wanted from respondents. The questionnaire was given to respondents who were stakeholders, namely supervisors, principals, teachers and parents of students who aimed to find out how the needs of these stakeholders were for the pedagogical literacy assessment model. The second test, according to Arikunto & Jabar (2004) in a book written by (Safithry, 2018) suggests that a tool or procedure used to determine or measure something using certain methods or rules.

Technical tests are conducted to develop products in the form of pedagogical literacy assessment models. The test is given to obtain the necessary data, after testing the test results, then validity is carried out using the Rasch model. Data collection techniques in the form of documentation that supports the research process and research results. The third researcher used documentation studies to examine the necessary documents related to the pedagogical literacy assessment model.

Data analysis was used to validate the pedagogical literacy assessment model instrument using the Rasch model, in R programming version 4.0.3 with the eRm package, namely on the validity of the content aspect construct, substantive aspect, structural aspect and external aspect. The steps taken to validate are as follows: (a) Data preparation in the form of inputting test takers' answers. The data used are all the answers of the test participants then typed into notepad. (b) validate the content aspect, (c) validate the substantive aspect. (c) validate the structural aspects. (d) validate the external aspects.

FINDINGS AND DISCUSSION

Construct Validity of Content Aspects of Teacher Pedagogical Literacy Assessment

Construct validity is validity that measures the extent to which test items are truly in accordance with the concept or conceptual definition that has been determined. Measurement of construct validity used the Rasch model on the questions on the pedagogical literacy assessment totalling 60 questions which were tested on 112 test participants. Construct validity measured includes content aspect construct validity, substance aspect construct validity. Rasch model-based construct validity on dichotomous items is done by analysing item fit to the model (Item fit). The content aspect construct validity test also has indicators in the form of item difficulty, respondent ability, item fit test, and respondent fit test. The results of the Item fit analysis of pedagogical literacy assessment items in table 1 below:

Table 1. Table of Non-Fit Items in the Item Fit Analysis of the Pedagogical Literacy Instrument for Teachers in the Junior High School Environment

No Butir	Chisq	df	p-Value	Outfit MSQ	Infit MSQ	Outfit t	Infit t
9	146.357	111	0.014	1.307	1.256	3.150	3.333
11	157.360	111	0.003	1.405	1.289	2.982	3.083
12	142.421	111	0.024	1.272	1.124	1.554	1.112
18	151.145	111	0.007	1.350	1.275	3.793	3.704
28	164.575	111	0.001	1.469	1.113	1.746	0.731
54	184.613	111	0.000	1.648	1.367	4.483	3.826
56	177.225	111	0.000	1.582	1.221	2.349	1.499

Item fit explains whether an item functions normally or not. Quantitatively, test items that are declared fit or can be said to behave consistently with what is expected by the first model are if the p-value is greater than the significance level ($p > 0.05$). From table 1, it can be seen that from 7 items out of 60 items of teacher pedagogical literacy assessment, there are items, namely questions 9, 11, 12, 18, 28, 54 and 56, that cannot be accepted as a suitable model with a 0.05 significance level. While the other 53 items can be accepted as questions that match the items with a significance level of 0.05. Meanwhile, when viewed from the parameter difficulty level of the item if the P value is as follows: $P \leq -2$ (items categorized as very easy), $-2 \leq P \leq -1$ (items categorized as easy), $-1 \leq P \leq 1$ (items categorized as medium), $1 \leq P \leq 2$ (items categorized as difficult), and $P \geq 2$ (items categorized as very difficult). The magnitude of the difficulty level value on each item can be seen in table 2.

Table 2 Level of Difficulty of the Items of the Instrument for Measuring Pedagogical Literacy of Teachers in the Junior High School Environment

Question	Score	Question	Score	Question	Score
1	-0.071	21	-0.280	41	-1.511
2	2.395	22	-0.542	42	0.092
3	0.255	23	-0.280	43	1.188
4	-0.875	24	0.295	44	-0.542
5	-1.809	25	-0.366	45	-1.031
6	-1.141	26	-1.198	46	-1.317
7	-0.030	27	-0.634	47	0.539
8	-1.444	28	2.085	48	0.214
9	0.580	29	0.580	49	-0.681
10	-1.085	30	0.133	50	0.789
11	1.051	31	0.874	51	0.052
12	1.485	32	-0.323	52	-0.728
13	0.174	33	-0.875	53	-0.196
14	-1.085	34	1.006	54	1.051
15	-2.859	35	0.376	55	0.918
16	0.874	36	-0.238	56	1.884
17	-0.154	37	0.498	57	-1.444
18	0.092	38	-0.978	58	0.457
19	-0.071	39	1.485	59	0.621
20	0.789	40	0.621	60	0.335

The table above shows the difference in the level of difficulty of each item. Based on the data output, it can be said that the level of difficulty of the items is within normal limits because it is still between the range of numbers -4 to 4. Based on table 4.28, there are 1 item classified as very easy, 13 items classified as easy, 39 items classified as medium, 5 items classified as difficult, 2 items classified as very difficult. The percentage of categories in the difficulty level of each item is: 1.67% of questions are categorized as very easy, 23.33% of questions are categorized as easy, 60% are categorized as moderate, 11.67% are categorized as difficult and 3.33% are categorized as very difficult. The difficulty level value ranges from -2.859 to 2.085. The question item with the lowest difficulty value is in question number 15 which is -2.859 and the highest difficulty value is in number 28 which reaches 2.085. From the table it can be seen that the level of item difficulty moves from -2.859 to 2.085. An effective test has an item difficulty level between -3.00 to 3.00.

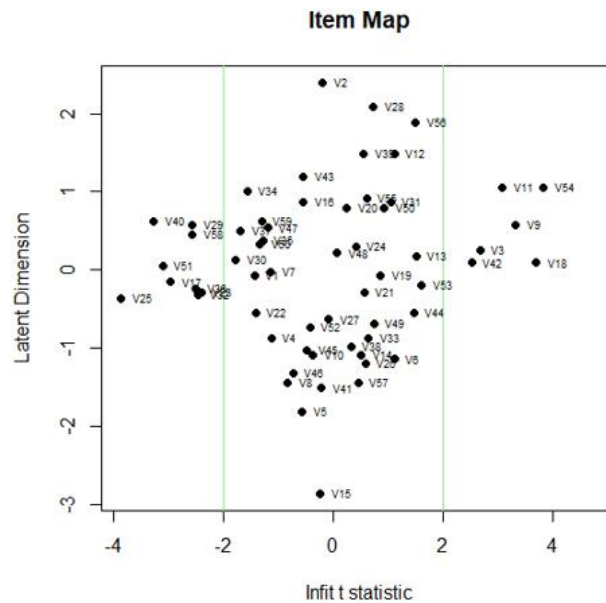


Figure 2. Item map of the measurement instrument for assessing pedagogical literacy of teachers in junior high schools

Figure 2 is a picture of the item map of the items of the pedagogical literacy measurement instrument for teachers with the help of eRm software. It can be seen that the item at the highest point is V2 which describes the most difficult item followed by number 28, and the lowest point is item number 15 which is the easiest item category.

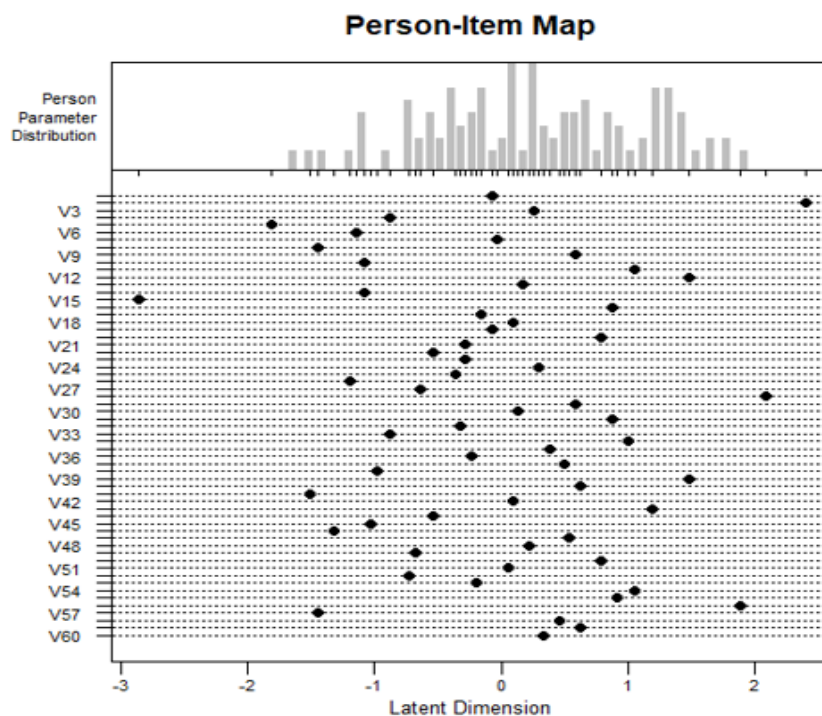


Figure 3 Person-Item Map of Instrument Items for Measuring Pedagogical Literacy of Teachers in the Junior High School Environment

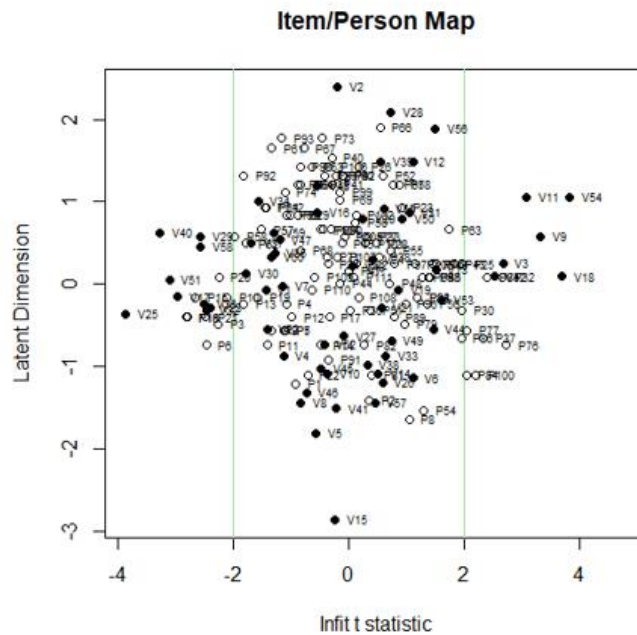


Figure 4 Item/Person Map of the Items of the Pedagogical Literacy Measurement Instrument in the Junior High School Environment

The evidence that the items of the pedagogical literacy measurement instrument are effectively used for test takers' abilities between -3.00 to 3.00 is explained by the item and test information functions in Figure 4. 4 explains that the information function will be maximized at the interval of learners' abilities between 0 to 3.0 and effective between -3.00 to 3.00.

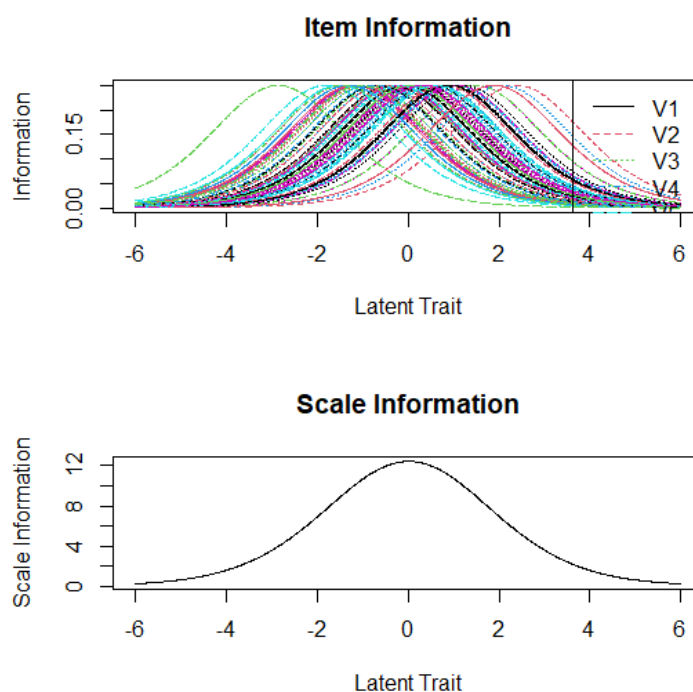


Figure 5. Information Function of the Pedagogical Literacy Assessment Instrument Items for Teachers in the Junior High School Environment

Construct Validity of Substantive Aspects of Pedagogical Literacy Assessment of Teachers in a Junior High School Environment



© 2025 The Author. This article is licensed CC BY SA 4.0.
visit [Creative Commons Attribution-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-sa/4.0/).



The quality of construct validity from the substantive aspect is used to test the test taker's ability to match the model. The substantive aspect is related to the content aspect. It is achieved by empirically finding to ensure that the test taker actually involves the ability of the measured field in answering the test items. Different response patterns are mismatches test takers who have a forgiveness (Θ) of 1.5 should be able to answer all items that have a difficulty level below 1.5, but in the field there are certainly some leftovers that are inconsistent or cause aberrant responses. How many test takers experience this aberrant response becomes substantive type construct validity.

This deviant response can be caused by inattention, cheating activities (cheating) or even misconceptions. Testing whether a person's response is deviant or not is called person fit. Quantitatively, the Outfit MSQ value response is between 0.5 to 1.5 while the outfit t value is between -2.0 to 2.0 and the chance of accepting H_0 (model fit) is greater than 0.05 ($p > 0.05$). Table 4.29 contains the results of the person fit test of 112 responses to the pedagogical literacy assessment test for teachers in the junior high school environment.

Table 3. Table of Test Takers with Aberrant Response

Peserta	Chisq	df	p-Value	Outfit MSQ	Infit MSQ	Outfit t	Infit t
P8	96.115	59	0.002	1.602	1.204	1.63	1.05
P32	95.433	59	0.002	1.591	1.294	3.43	2.67
P54	91.655	59	0.004	1.528	1.245	1.56	1.30
P84	118.371	59	0.000	1.973	1.320	3.25	2.04
P85	88.190	59	0.008	1.470	1.263	2.81	2.41
P100	124.404	59	0.000	2.073	1.349	3.51	2.20

Of the 112 test takers there were 6 test takers who experienced responses that deviated from the model. It can be seen from the six test takers that three test takers (P32, P84, P100) did not meet all person fit criteria. Two test takers did not meet the p-value criteria ($p < 0.05$) and the MSQ outfit (0.5 to 1.5) there were 2 test takers (P8, P54). While one participant (P85) also did not meet the two criteria but at p-value ($p < 0.05$) and infit t (-2.0 to 2.0).

From the explanation in the table it can be concluded that there are 94.64% of test takers' responses that are reasonable according to the model or do not experience deviation (Aberrant Response), while 5.36% of test takers' responses experience deviation. The large percentage of test takers who have reasonable responses according to the model can be used as a basis that the test sufficiently meets substantive validity. Even if we use a significance level of 0.01, only two test takers did not fit the model.

Construct Validity of Structural Aspects of the Pedagogical Literacy Assessment of Teachers in a Junior High School Environment

Indicators of teacher pedagogical literacy assessment have structural aspect construct validity that is the test is unidimensional. A teacher pedagogical literacy test built on a one-dimensional framework must be valid to have meaning. The principle of testing is first stated using the null hypothesis which states that the second eigenvalue is greater than the first eigenvalue.

Test Results of the Instrument Items of the Pedagogical Literacy Measurement Instrument for Teachers in the Junior High School Environment shows:

Alternative hypothesis: the second eigenvalue of the observed data is substantially larger than second eigenvalue of data under the assumed IRT model. Second eigenvalue in the observed data: 0.2254. Average of second eigenvalues in Monte Carlo samples: 0.2343

Furthermore, the Kaiser-Meyer-Olkin factor adequacy test was conducted on the 60 pedagogical literacy questions used to assess the adequacy of the sample for factor analysis. KMO measures how well the data are suitable for factor analysis by assessing the proportion of variance among variables that may be common variance. The result was an MSA (Measure of Sampling Adequacy) value of 0.57. The value obtained above the 0.50 threshold indicates that the data has a good enough correlation between variables, so factor analysis can be continued. The results of the unidimensionality test in the context of factor analysis can also be seen in the scree plot which shows the distribution of eigenvalues of the extracted factors. The following are the results of the scree plot of the Pedagogical Literacy Assessment of Teachers in the Junior High School Environment:

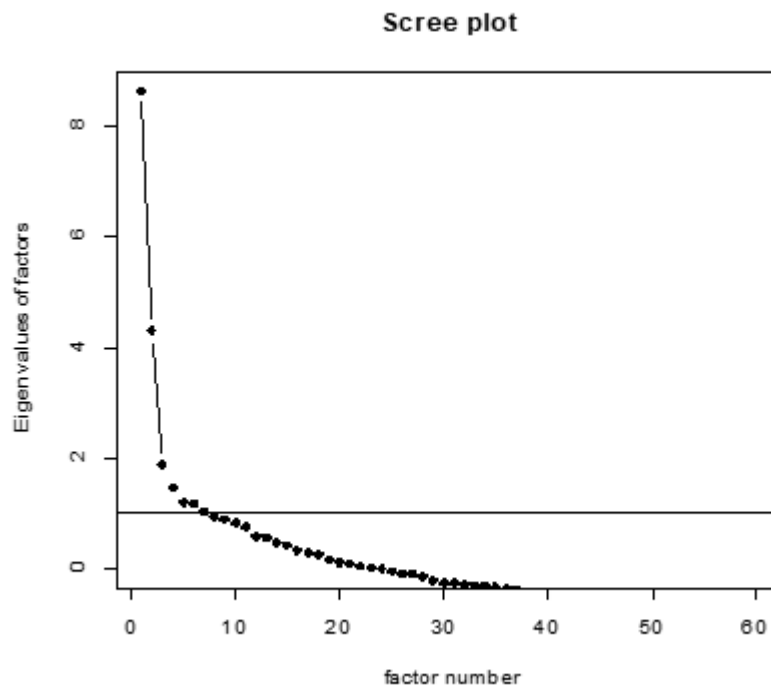


Figure 6. Scree plot of pedagogical literacy assessment items in the junior high school environment

The scree plot shows that the items on the pedagogical literacy assessment in the middle school environment are assumed to have one dimension, it can be seen that there is one dominant factor that has a higher eigenvalue than the other factors. So based on the unidimensional feasibility criteria, the Pedagogical Literacy Assessment of Teachers in the Middle School Environment is fulfilled. Because there are three eligibility criteria that are met, first, the second eigenvalue of 0.2343 is greater than the first eigenvalue of 0.2254. Second, having a KMO value of 0.57 is greater than the threshold limit of 0.5 and third, there is a dominant factor that has a higher eigenvalue than other factors.

Construct Validity of External Aspects of Pedagogical Literacy Assessment of Teachers in a Junior High School Environment

The assessment of pedagogical literacy from the external aspect of construct validity is carried out to ensure that the research results can be generalized and applied to a wider context outside the sample studied. To determine the construct validity of external aspects, one of the approaches used is Person Separation Reliability or person separation. This person separation is used to enable researchers to understand the effectiveness of measurement instruments. By conducting this analysis, areas of improvement can be identified to increase the accuracy and reliability of the measuring instrument, as well as ensuring that the

instrument can properly distinguish between individuals with different levels of ability. The results of the person separation analysis using eRm can be seen in the following table:

Person separation Reliabilty of Pedagogical Literacy Assessment Instrument Items in the Junior High School Environment

Separation Reliability : 0.8645
Observed Variance : 0.6667 (Squared Standard Deviation)
Mean Square easurement Error: 0.0904 (Model error Variance)

From the table it can be seen that the Person Separation Reliability value is 0.8645. According to (Salsabila, 2023) the person separation value can be determined as follows:

$$H = \frac{(4 \times \text{Separation}) + 1}{3}$$

Thus the person separation value for the test is 1.486. From the person separation value, it can be seen that the classification of test takers can only be 1 (rounding from 1.486). This means that the instrument that has been made only distinguishes test takers in two categories, namely complete and incomplete. The consequence is that the results of the teacher pedagogical literacy assessment only distinguish test takers into two groups, namely test takers who have had minimal adequacy and have not had minimal adequacy of teacher pedagogical literacy skills.

CONCLUSIONS

The test construction of the Pedagogical Literacy Assessment Model for Teachers in the Junior High School Environment in the form of multiple choice 1) At the 95% confidence level there are 53 items that match the modeling, (2) The level of item difficulty is in the range of -3 to 3, meaning that the test items are suitable for the test takers' abilities. (3) At the 95% confidence level, 94.64% of the test takers' responses matched the modeling. (4) The pedagogical literacy model assessment test instrument in the junior high school environment has the eligibility criteria of the unidimensional concept that each item only measures one ability. (5) The pedagogical literacy model assessment instrument in the Junior High School environment distinguishes test takers into two categories, namely test takers who have minimal adequacy and have not experienced minimal adequacy.

REFERENCES

- Agung Santoso, "Rumus Slovin: Panacea Masalah Ukuran Sampel?," J. Psikol. Univ. Sanata Dharma, vol. 4, 2023, [Online]. Available: <https://e-journal.usd.ac.id/index.php/suksma/article/view/6434>
- D. Pratama, "Analisis Kualitas Tes Buatan Guru Melalui Pendekatan Item Response Theory (IRT) Model Rasch," Tarbawy J. Pendidik. Islam, vol. 7, no. 1, pp. 61-70, 2020, doi: 10.32923/tarbawy.v7i1.1187.
- Djaali and Pudji Muljono, "Pengukuran Dalam Bidang pendidikan.pdf."
- Esty Aryani Safithry, asesmen teknik tes dan non tes. Malang: CV IRDH, 2018. [Online]. Available: https://www.google.co.id/books/edition/ASESMEN_TEKNIK_TES_DAN_NON_TES/EcbODwAAQBAJ?hl=id&gbpv=1&dq=asesmen+teknik+tes+dan+non+tes&prints=frontcover
- F. Hidayat and M. Nizar, "Model Addie (Analysis, Design, Development, Implementation and Evaluation) Dalam Pembelajaran Pendidikan Agama Islam," J. Inov. Pendidik. Agama Islam, vol. 1, no. 1, pp. 28-38, 2021, doi: 10.15575/jipai.v1i1.11042.
- Indri Pudjiati and Faisal Madani, "Asesmen Autentik Analisis Butir Soal Dengan Rasch

- Model Di Sekolah Dasar: Literature Review,” Mutiara J. Penelit. dan Karya Ilm., vol. 1, no. 4, pp. 01–12, 2023, doi: 10.59059/mutiara.v1i4.325.
- M. R. S. S. N. Ayub, E. Istiyono, S. Munadi, C. Permadi, A. Pattiserlihun, and D. N. Sudjito, “Analisa Penilaian Soal Fisika Menggunakan Model Rasch Dengan Program R,” J. Sains dan Edukasi Sains, vol. 3, no. 2, pp. 46–52, 2020, doi: 10.24246/juses.v3i2p46-52.
- M. Safitri and M. R. Aziz, “ADDIE, sebuah model untuk pengembangan multimedia learning,” J. Pendidik. Dasar, vol. 3, no. 2, pp. 50–58, 2022, [Online]. Available: <http://jurnal.umpwr.ac.id/index.php/jpd/article/view/2237>
- Okpatrioka, “Research And Development (R & D) Penelitian yang Inovatif dalam Pendidikan,” J. Pendidikan, Bhs. dan Budaya, vol. 1, no. 1, pp. 86–100, 2023.
- S. Mulyani, R. Efendi, and T. R. Ramalis, “Karakterisasi Tes Keterampilan Pemecahan Masalah Fisika Berdasarkan Teori Respon Butir,” J. Pendidik. dan Ilmu Fis., vol. 1, no. 1, p. 1, 2021, doi: 10.52434/jpif.v1i1.1006.
- W. Lestari, I. Wigati, M. I. Sholeh, D. Desi Pramita, U. Islam Negeri Raden Fatah Palembang, and S. Selatan, “Instrumen Literasi Digital Guru Menggunakan Model Rasch,” Orbital J. Pendidik. Kim., vol. 6, no. 2, 2022.
- Yudi Hari Rayanto & Sugianti, Penelitian Pengembangan Model Addie Dan R2D2, vol. 216. Pasuruan: Lembaga Academic & Research Institut, 2020. [Online]. Available: https://www.google.co.id/books/edition/PENELITIAN_PENGEMBANGAN_MODEL_ADDIE_DAN/pJHcDwAAQBAJ?hl=id&gbpv=1&dq=penelitian+model+pengembangan&printsec=frontcover