

## Analysis of the Validity of the Instrument for Early Childhood Numeracy Literacy Assessment

Rivo Panji Yudha, Teguh Samudra Hetharia

Master's Program in Early Childhood Education, Panca Sakti University Bekasi, Indonesia

### ABSTRACT

Numeracy literacy in early childhood is a fundamental aspect of cognitive development; however, valid assessment instruments remain limited in Indonesia. This research aims to analyze the validity of numeracy literacy assessment instruments for early childhood in Cakranegara District, Mataram, Lombok. Employing a descriptive quantitative approach, the study involved 135 children with data collection methods through direct observation and numeracy tests. Instrument validation was conducted through content validity testing by three early childhood education experts and exploratory factor analysis (EFA) using Principal Component Analysis method. The research results indicate that the prototype instrument achieved an average validation score of 87.11%, with a Kaiser-Meyer-Olkin (KMO) value of 0.725 and a significant Bartlett's test ( $\chi^2 = 968.515$ ;  $p < 0.000$ ). The inter-rater reliability test yielded an average value of 0.85, falling into the high category. This research produces valid, reliable, and contextual numeracy literacy assessment instruments that can be used to identify the numeracy development of early childhood children more accurately and effectively.

**Keywords:** Numeracy Literacy, Early Childhood Assessment, Instrument Validity

#### Corresponding author

**Name:** as Rivo Panji Yudha

**Email:** rivoyudha@yahoo.co.id

### INTRODUCTION

Numeracy literacy in early childhood is a fundamental aspect of cognitive development that contributes to school readiness and academic achievement at subsequent levels (Tazouti et al., 2024). According to the National Council of Teachers of Mathematics (Lindquist & Gates, 2020), valid and reliable numeracy literacy assessments are necessary to ensure that children acquire a strong mathematical foundation from an early age. The education curriculum in Indonesia also emphasizes the importance of numeracy literacy from early childhood as stated in the Minister of Education and Culture Regulation (Permendikbud) No. 137 of 2014 concerning National Standards for Early Childhood Education. Based on the child development theory proposed by Piaget, young children are in the pre-operational stage, which requires concrete and contextual numeracy

stimulation, necessitating assessments that are appropriate for their developmental stage (Piaget, 2013).

However, in the field, many literacy and numeracy assessment instruments used in early childhood education institutions have not been standardized in a valid and reliable manner. Some measurement tools used by educators are not aligned with children's cognitive development, leading to bias in assessment results (Dong et al., 2021; Hellstrand et al., 2020). A study by Aprilia et al. (2023) reveals that in various regions of Indonesia, the use of assessment instruments remains subjective and lacks strong scientific evidence (Aprilia et al., 2023). Furthermore, research by Wahyuni & Aprilia et al. (2022) found that the majority of early childhood educators in Indonesia still face challenges in understanding and implementing literacy and numeracy assessments that meet international standards (Oktafiani et al., 2022).

The absence of valid assessment instruments can hinder the accurate evaluation of numeracy development in early childhood. In the context of education in Cakranegara District, Mataram, Lombok, preliminary studies indicate that the majority of early childhood education units still utilize non-standard assessments, which may lead to less accurate diagnoses of children's numeracy skills. This can result in less effective interventions in enhancing early numeracy literacy. Therefore, this research is urgent in developing and analyzing the validity of numeracy literacy assessment instruments that can be used as a standard in assessing early childhood development.

Several previous studies have examined the importance of numeracy literacy in early childhood (Nur et al., 2022; Purpura & Lonigan, 2015; Raghubar & Barnes, 2017; Rahmadeni, 2022); however, research on the validity of numeracy literacy assessment instruments in the educational context of Indonesia, particularly in the Cakranegara area, remains limited. This study aims to fill the research gap by developing assessment instruments that have been empirically validated in accordance with the local context and characteristics of early childhood in Indonesia.

The phenomenon in the field indicates a gap between the national education policy that mandates standardized literacy and numeracy assessments and the implementation occurring in early childhood education units. Many educators still rely on observation-based assessments without standardized measurement tools (Fatmawati & Aziz, 2022; Rahman, 2019; Suyadi, 2016). Furthermore, most research on literacy and numeracy assessments has focused more on primary and secondary education levels, while numeracy assessments in early childhood education have received less attention.

The novelty of this research lies in the development of an assessment instrument that is empirically validated and contextually based, particularly in the early childhood education environment in Cakranegara District, Mataram, Lombok. Unlike previous studies that emphasized assessment theory, this research combines instrument validation through empirical testing and adaptation to local contexts, making it more applicable for educators in Indonesia.

Based on this background, this study aims to analyze the validity of the early childhood numeracy literacy assessment instrument in Cakranegara District, Mataram,

Lombok, so that it can be used as an accurate measurement tool in identifying early childhood numeracy development precisely and effectively.

## METHOD

This research employs a quantitative approach with a descriptive research type. Descriptive research aims to provide a clear and systematic depiction of a particular phenomenon or condition. In this context, this study seeks to analyze the validity of the assessment instruments for early childhood numeracy literacy, without any intervention or special treatment. The population in this study consists of all early childhood children attending school in the Cakranegara District, Mataram, Lombok. The sample was selected using purposive sampling techniques, with the criteria of children who have received basic numeracy education in early childhood education institutions. The research instrument consists of a numeracy literacy test that has been developed and validated through content validity testing using expert judgment methods and construct validity using exploratory factor analysis (EFA). Data collection procedures were carried out through direct observation, as well as numeracy tests administered to the children who served as research respondents. To test the reliability of the evaluation instrument for the theoretical trial stage of the internship, inter-rater reliability is used, which is the reliability observed from the level of agreement between raters. Inter-rater reliability (IRR) will provide an overview (in the form of scores) of the extent of consensus or agreement given by experts. The IRR coefficient used is Cohen's Kappa (K) agreement coefficient. The measurement tool has adequate stability if the IRR between measurements is  $> 0.50$ , and high stability if the IRR between measurements is  $\geq 0.61$  (Streiner & Norman, 2008).

**Table 1. Criteria for Decision-Making Based on Expert Judgment**

Category	Validity Level	Criteria
81 -100	Very High	It can be used without revision.
61 - 80	High	It can be used with minor revisions.
41 - 60	Considerable	It can be used with major revisions.
21 - 40	Low	Review a lot and revalidate.
0 - 20	Very Low	The instrument cannot be used.

(Yudha, 2020)

This method was chosen because it aligns with the research objectives to develop and test the validity of the numeracy literacy assessment instrument. To ensure the validity of the research results, a trial of the instrument was conducted on a small group before being applied to a larger population. The reliability of the instrument was tested using the test-retest method to measure the consistency of assessment results over different times. Thus, this research is expected to produce a valid and reliable numeracy literacy assessment instrument that can be applied in the context of early childhood education in Indonesia.

## FINDING AND DISCUSSION

### RESULT

The expert validation stage is an important part of the assessment instrument development process, particularly to ensure that the resulting product has a high level of accuracy, relevance, and utility. In the development of numeracy literacy assessment instruments for early childhood, expert validation aims to evaluate the feasibility of the instrument in terms of content, construct, language, and practicality. This validation involves three experts with expertise in early childhood education, assessment, and learning practices.

Through expert validation, the instrument is assessed based on indicators designed in accordance with the research objectives and field needs. This process not only provides a quantitative assessment in the form of scores but also qualitative input in the form of suggestions and constructive criticism for product improvement. After validation, revisions are made to refine the instrument before it is tested on the main respondents, namely children in early childhood education institutions who are the subjects of the research.

**Table 2. Expert Validation Results**

Assessment Aspects	Assessment Indicators	Validator			Total	Presentation		
		1	2	3		1	2	3
Content Eligibility	Conformity of content to learning objectives	5	4	4	13			
	Suitability of content to the characteristics of early childhood	5	4	5	14			
	Conformity of content with scientific literacy indicators	4	5	5	14	88	92	92
	Conformity of content with an observation-based approach	4	5	4	13			
	Accuracy of the concepts conveyed	4	5	5	14			
	Clarity of the language used	4	4	4	12			
Language Qualification	The simplicity of the language corresponds to the level of understanding of young children	4	4	5	13	80	85	95
	Consistency of terms used	4	5	5	14			
	Suitability of language to local context	4	4	5	13			
Visual Qualifications	Suitability of visual design to the characteristics of early childhood	5	4	4	13			
	Text readability (size, font type, and color)	4	5	4	13	93	87	87
Practicality of Use	Layout suitability	5	4	5	14			
	Ease of use of the instrument by the user	4	5	4	13			
	Time required to use the instrument	4	4	4	12	80	80	87
	Availability of resources required for use	4	3	5	12			
<b>Average</b>					<b>87.11</b>			

The initial prototype was evaluated by three experts through a validation form that included aspects of content, language, visuals, and practicality. The assessment experts determined that the indicators and rubrics were aligned with the developed science literacy domain. Recommendations from these experts included simplifying the language to match the understanding level of early childhood education teachers.

The assessment experts and learning practitioners provided a high score of 92% for the content relevance and language appropriateness. Overall, the questions are quite relevant, but some questions related to geometric shapes could be more challenging. For instance, instead of merely showing an image and asking the child to guess the shape, I suggest providing a variety of questions. Regarding language appropriateness, the validators also indicated that language simplification should be carried out to ensure that the question instructions are easily understood by children. For example, the sentence 'Count how many of these objects' could be more specific.

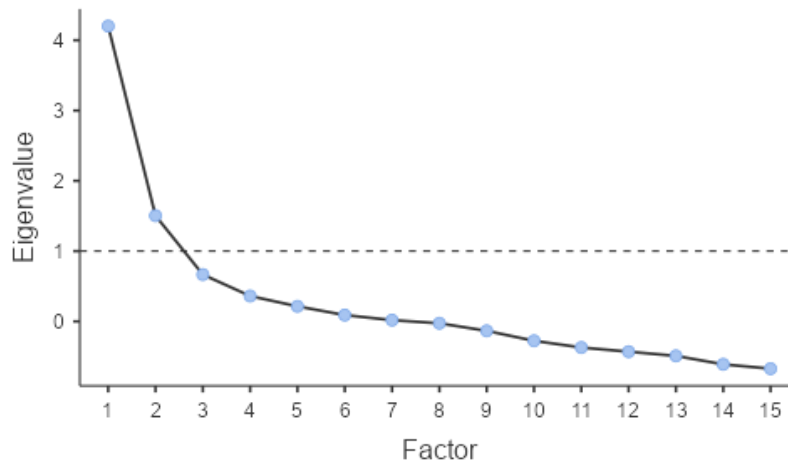
Early Childhood Education Experts gave a high score of 92% for visual suitability with 93%. Overall, the experts recommend that this instrument should include more images that reflect objects identifiable by children, such as pictures of animals or items from their daily lives. The design phase produces an initial prototype of a valid assessment model, with an average validation score of 87.11% (very valid category). Expert validation indicates that this assessment instrument is suitable for testing, with several revisions to enhance practicality and appeal.

Exploratory Factor Analysis was performed from data collected from 135 children using the Primary Component Analysis (PCA) and the Kaiser-Meyer-Olkin and Bartlett Tests were applied before evaluating the results of the exploratory factor analysis with the Kaiser Normalization value of the KMO Index ranging from 0 to 1, with 0.60 considered suitable for factor analysis (Ebneter & Latner, 2013; Williams et al., 2012).

**Table 3. KMO (Kaiser-Meyer-Olkin) Test**

<b>KMO and Bartlett's Test</b>		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.725
Bartlett's Test of Sphericity	Approx. Chi-Square	968.515
	df	105
	Sig.	.000

The KMO result in this study is 0.725, indicating that the sample is adequate for conducting factor analysis, and Bartlett's Test of Sphericity was significant ( $\chi^2 = 968.515$ ;  $df = 105$ ,  $p < 0.000$ ; Table 3) showing that the relationships between variables are strong and the data is suitable for conducting Exploratory Factor Analysis. A probability value  $< 0.05$  indicates that the sample size and normality are adequate for performing PCA. The item loading presented in Table 4 indicates that items are recommended to be removed from the scale if they do not meet the criterion of a coefficient value lower than 0.50.



**Gambar 1. Scree Plot**

Figure 1 is a representation of the total variance explained table, observing the condition of the Y-axis which has an eigenvalue >1 (Faller et al., 2006). In the scree plot above, there are six factors with an eigenvalue >1, namely the first factor, the second factor, the third factor, and the fourth factor.

The next stage is to test the reliability of the instrument. The following are the results of the reliability test simulation of the instrument using Inter Rater Reliability (IRR) analyzed with SPSS version 28. This data illustrates the reliability among raters regarding the assessment instrument for children's numeracy literacy.

**Table 4. Results of Reliability Testing**

Aspect	Number of Observation Items	Inter Rater Reliability Value	Reliability Category
Content Eligibility	4	0,87	High
Language Qualification	4	0,85	High
Visual Qualifications	3	0,83	High
Practicality	3	0,84	High
<b>Average</b>	<b>14</b>	<b>0,85</b>	<b>High</b>

From the output table above, it is found that the average reliability value is 0.85, which is in the high category. This shows that the assessment between raters of the instrument items is quite consistent. The visual suitability aspect has a lower IRR value than other aspects, namely 0.83, but is still in the high category. This can be caused by variations in interpretation between raters regarding validator responses which tend to be more subjective in the visual aspect.

## DISCUSSION

The findings of this research contribute significantly to the development of assessment instruments for numeracy literacy for early childhood in the Cakranegara region, Lombok. The expert validation process resulted in a prototype instrument with an average score of 87.11%, which is categorized as highly valid. This indicates that the developed instrument has met the required criteria of content, language, visual, and practicality in early childhood education assessment. Exploratory Factor Analysis reveals interesting results, with a Kaiser-Meyer-Olkin (KMO) value of 0.725, indicating the adequacy of the sample for factor analysis. Bartlett's test shows strong significance ( $\chi^2 = 968.515$ ;  $p < 0.000$ ), indicating a strong relationship among the variables. Furthermore, the scree plot analysis identifies six factors with eigenvalues greater than 1, providing an empirical basis for the instrument's structure. The inter-rater reliability test yielded an average score of 0.85, which falls into the high category. This finding is consistent with previous research by Dong et al. (2021) and Hellstrand et al. (2020), which emphasizes the importance of valid and reliable assessment instruments in measuring early childhood numeracy skills. Nevertheless, the visual aspects show slight variations in interpretation among raters, which may be attributed to the subjective nature of assessing visual designs.

The unique contribution of this research lies in its approach that combines empirical validation with local context, which differs from previous studies that focused more on assessment theory. This is highly relevant in the context of education in Indonesia, where there is still a gap between national policy and practical implementation in early childhood numeracy literacy assessment. The practical implications of this research are crucial for educators and early childhood education practitioners. The developed instrument can assist them in conducting assessments that are more accurate, objective, and aligned with the cognitive development stages of children. However, this research has limitations, particularly in its geographical scope limited to the Cakranegara area of Lombok, thus generalization of the results should be done with caution.

For future research, it is recommended to expand the sample scope, develop instruments with more complex question variations, and conduct trials in various educational contexts. Focusing on simplifying language and using more contextual images may also be promising areas for the development of early childhood numeracy literacy assessment instruments.

## CONCLUSION

This research successfully developed and validated a numeracy literacy assessment instrument for early childhood in the Cakranegara area, Mataram, Lombok, which holds methodological and practical significance in the context of Indonesian education. Through a rigorous validation process, including expert assessments and exploratory factor analysis, the developed instrument demonstrates a very high level of validity quality. This indicates that the instrument not only meets scientific standards but can also be effectively applied in the context of early childhood education.

The main contribution of this research lies in the development of an empirical and contextual assessment approach. Unlike previous instruments that tended to be subjective, this study produces a measurement tool with high inter-rater reliability, allowing for a more objective and accurate evaluation of children's numeracy skills. The exploratory factor analysis conducted reveals the complex structure of the instrument with six main factors, providing a strong scientific foundation for the assessment of early childhood numeracy skills.

The findings of this research have important implications for early childhood education practices in Indonesia. First, this instrument can help address the gap between national education policy and practical implementation in literacy and numeracy assessment. Second, the developed approach can serve as a model for early childhood education institutions to design assessment instruments that are more systematic, valid, and aligned with children's cognitive development stages. Nevertheless, this research is not without limitations. The limited geographical scope to the Cakranegara area in Lombok necessitates caution in generalizing the results. Therefore, further research is highly recommended to expand the sample scope, develop more complex question variations, and conduct trials in various educational contexts.

The practical recommendation from this research is the necessity for the systematic adoption of assessment instruments that are scientifically validated by early childhood education institutions. Educators and education practitioners are advised not only to utilize these instruments but also to continuously develop and adapt them according to local characteristics and contemporary developments in the field of early childhood education. The final conclusion of this research emphasizes that the development of valid, reliable, and contextual numeracy literacy assessment instruments is a strategic step in enhancing the quality of early childhood education. With a systematic and evidence-based approach, we can more effectively understand, assess, and support the development of children's numeracy skills from an early age.

## REFERENCES

- Aprilia, N., Setiani, Y., & Hadi FS, C. A. (2023). Pengembangan Instrumen Tes Numerasi Pada Asesmen Kompetensi Minimum Yang Bernilai Budaya Lokal. *Jurnal Educatio FKIP UNMA*. <https://doi.org/10.31949/educatio.v9i2.4824>
- Dong, Y., Clements, D. H., Day-Hess, C. A., Sarama, J., & Dumas, D. (2021). Measuring Early Childhood Mathematical Cognition: Validating and Equating Two Forms of the Research-Based Early Mathematics Assessment. *Journal of Psychoeducational Assessment*. <https://doi.org/10.1177/07342829211037195>
- Ebneter, D. S., & Latner, J. D. (2013). Stigmatizing attitudes differ across mental health disorders: A comparison of stigma across eating disorders, obesity, and major depressive disorder. *Journal of Nervous and Mental Disease*. <https://doi.org/10.1097/NMD.0b013e318288e23f>
- Faller, H., Kohlmann, T., Zwingmann, C., & Maurischat, C. (2006). Exploratory and confirmatory factor analysis. *Rehabilitation*. <https://doi.org/10.1055/s-2006->

940029

- Fatmawati, D. S., & Aziz, H. (2022). Studi Analisis Pelaksanaan Asesmen terhadap Perkembangan Anak Usia Dini di KB X Pangandaran. *Jurnal Riset Pendidikan Guru Paud*. <https://doi.org/10.29313/jrpgp.v1i2.532>
- Hellstrand, H., Korhonen, J., Räsänen, P., Linnanmäki, K., & Aunio, P. (2020). Reliability and validity evidence of the early numeracy test for identifying children at risk for mathematical learning difficulties. *International Journal of Educational Research*. <https://doi.org/10.1016/j.ijer.2020.101580>
- Lindquist, M. M., & Gates, J. D. (2020). National Council of Teachers of Mathematics. *Teaching Children Mathematics*. <https://doi.org/10.5951/tcm.1.1.0053>
- Nur, I. R. D., Herman, T., & Dahlan, T. H. (2022). Numeracy Literacy in Early Childhood: An Investigation in Arithmetic, Geometry and Patterns in Early Stage. *JTAM (Jurnal Teori Dan Aplikasi Matematika)*. <https://doi.org/10.31764/jtam.v6i2.7007>
- Oktafiani, N., Witono, A. H., & Ermiana, I. (2022). Kesiapan Guru dalam Melaksanakan Kegiatan Literasi dan Numerasi di SDN 1 Selebung. *Jurnal Ilmiah Profesi Pendidikan*. <https://doi.org/10.29303/jipp.v7i4b.1049>
- Piaget, J. (2013). Origin of Intelligence in the Child. In *Origin of Intelligence in the Child*. <https://doi.org/10.4324/9781315006260>
- Purpura, D. J., & Lonigan, C. J. (2015). Early Numeracy Assessment: The Development of the Preschool Early Numeracy Scales. *Early Education and Development*. <https://doi.org/10.1080/10409289.2015.991084>
- Raghubar, K. P., & Barnes, M. A. (2017). Early numeracy skills in preschool-aged children: a review of neurocognitive findings and implications for assessment and intervention. In *Clinical Neuropsychologist*. <https://doi.org/10.1080/13854046.2016.1259387>
- Rahmadeni, F. (2022). Urgensi Pengenalan Konsep Literasi Numerasi pada Anak Usia Dini. *ARITHMETIC: Academic Journal of Math*. <https://doi.org/10.29240/ja.v4i1.4626>
- Rahman, M. H. (2019). TEKNIK ASSESMEN OBSERVASI DAN PERCAKAPAN SEBAGAI UPAYA PEMANTAUAN PERKEMBANGAN BAHASA ANAK DI TK PAMARDI SIWI NANGGULAN MAGUWO HARJO SLEMAN YOGYAKARTA. *Tumbuh Kembang: Kajian Teori Dan Pembelajaran PAUD*. <https://doi.org/10.36706/jtk.v6i1.8351>
- Streiner, D. L., & Norman, G. R. (2008). Health Measurement Scales: A practical guide to their development and use. In *Health Measurement Scales: A Practical Guide to their Development and Use*. <https://doi.org/10.1093/acprof:oso/9780199231881.001.0001>
- Suyadi. (2016). Perencanaan dan Asesmen Perkembangan pada Anak Usia Dini (Studi Kasus pada Lembaga Pendidikan Anak Usia Dini (PAUD/TK/RA) di Daerah Istimewa Yogyakarta). *Golden Age: Jurnal Ilmiah Tumbuh Kembang Anak Usia Dini*.
- Tazouti, Y., Thomas, A., Hoareau, L., Jarlégan, A., Hubert, B., & Luxembourger, C. (2024). Assessment of an educational classroom app's impact on preschoolers' early numeracy skills. *European Journal of Psychology of Education*. <https://doi.org/10.1007/s10212-023-00698-1>

- Williams, B., Brown, T., & Onsman, A. (2012). Exploratory factor analysis : A five-step guide for novices. *Journal of Emergency Primary Health Care (JEPHC)*. <https://doi.org/10.1080/09585190701763982>
- Yudha, R. P. (2020). Validity And Reliability Rubric Of Performance Assessment Geometry Study In Junior High School Using The Many Facet Rasch Model Approach. *Eduma : Mathematics Education Learning and Teaching*. <https://doi.org/10.24235/eduma.v9i2.7100>