International Journal of Teaching and Learning (INJOTEL) e-ISSN: 3025-8308
Vol. 3 No. 11 2026, pages: 250~272

MODEL ASSESSMENT OF SCIENTIFIC LITERACY BASED ON OBSERVATION OF CHILDREN AGED 5-6 YEARS: DEVELOPMENT STUDY

Rivo Panji Yudha
Master's Program in Early Childhood Education, Panca Sakti University Bekasi, Indonesia
Corresponding author email: rivoyudha@yahoo.co.id

Siti Aisyah
Master's Program in Early Childhood Education, Panca Sakti University Bekasi, Indonesia
trixie.fzaneta@gmail.com

Keywords

Assessment Model, Science Literacy, Observation, Early Childhood.

Abstract

This research aims to determine the development of an observation-based science literacy assessment model for early childhood in Padang Panjang Village, to assess the feasibility of the observation-based science literacy assessment model for early childhood in Padang Panjang Village as valid and reliable, and to evaluate the science literacy skills of children before and after using the observation-based science literacy assessment model. This research is conducted in all kindergartens in Padang Panjang Village, Tanta District, Tabalong Regency, South Kalimantan Province. This research is a type of research and development. The research and development model applied is the 4D (Four-D) development model, which includes the stages of definition research, design, and development. The research results indicate that 1) the initial stage of the study reveals a substantial gap in the practice of science literacy assessment at the early childhood education level. As many as 70% of teachers acknowledge that they do not possess systematic assessment instruments, indicating methodological weaknesses in measuring children's science abilities. 2) the feasibility of the assessment model instruments: These assessment instruments are deemed suitable for use, with a very high validity level (86.02%) and strong reliability (average ICC value of 0.84).

INTRODUCTION

Scientific literacy in early childhood is an important foundation for developing scientific thinking skills and fundamental exploration abilities essential for their intellectual development. Ideally, scientific literacy is not merely an understanding of simple scientific concepts, but rather the child's ability to observe, question, explore, and build an initial understanding of the natural phenomena around them. Recent research indicates that stimulating scientific literacy from an early age can significantly contribute to children's cognitive development, creativity, and critical thinking skills.

Early childhood children need to acquire science literacy (Fang & Wei, 2010). Early science experiences from preschool to third grade play a crucial role in enhancing children's knowledge, skills, and dispositions (Mantzicopoulos et al., 2013) that are necessary for future employment and preparing students for an economy that demands innovative solutions to complex problems (Park et al., 2017) argue that science education (Aronin & Floyd, 2013) should begin in early childhood because the concepts taught foster curiosity, creativity, collaboration, and critical thinking that children need.

Although the increasing need for children (ages 3-8) to begin science education in preschool has been justified, the inadequacy of science education in early childhood causes early childhood education teachers to avoid teaching science and thus fail to develop their confidence in teaching related subjects in the classroom (Dunlosky et al., 2013; Fenty & Anderson, 2014). Teachers' beliefs about teaching have been studied from various perspectives, including how teachers' beliefs influence (a) decision-making and instructional practices, (b) classroom interpretation and practices.

Various international studies, such as the research conducted by Harlen & Qualter (2018) in the book "The Teaching of Science in Primary Schools," emphasize that science literacy in early childhood plays a strategic role in building the foundation of scientific abilities. They reveal that children who receive early science stimulation have better observation, classification, and inference skills compared to children who do not have similar experiences (Harlen & Qualter, 2018). Furthermore, the research by Kallery & Psillos indicates that observational approaches can enhance children's science literacy by up to 65% more effectively than conventional methods (Kallery & Psillos, 2002).

Previous research has demonstrated the importance of developing a child development-based science literacy assessment model. The study conducted by Akerson et al. (2018) revealed that observation-based assessments are more suitable for early childhood than written tests (Akerson et al., 2018). This assessment allows for the observation of children's abilities through play and exploration activities. Similarly, previous research emphasizes that measuring the science abilities of young children should focus on the processes of exploration, observation, and active engagement of the child (Pereira et al., 2020; Si et al., 2017).

However, the reality on the ground shows a different condition. Based on the results of the pre-observation conducted at the kindergarten in Padang Panjang Village, Tanta District, Tabalong Regency, several fundamental issues were found. Initial interviews with the principal and several teachers revealed that: (1) 70% of teachers do not yet have comprehensive and systematic science literacy assessment instruments, (2) the learning approach is still conventional with lecture methods and task assignments, (3) there is no specific assessment model to measure the science literacy of children aged 5-6 years.

There is a significant gap between the ideal of science literacy and the reality of its implementation in the field. Previous research has predominantly focused on the concept of science literacy in general, yet has not produced a specific, comprehensive, and observation-based assessment model for children aged 5-6 years. Earlier studies have also tended to concentrate more on the development of science learning models for early childhood rather than on the development of assessment models. Conversely, studies related to observation-based science literacy assessment remain very limited, especially those designed for early childhood in rural areas such as Padang Panjang Village. Meanwhile, the observable phenomenon gap is the misalignment between children's potential abilities to explore scientific phenomena and the limitations of the assessment instruments available to educators. The gap phenomenon is also evident from the field conditions. Most kindergartens in the area do not yet have systematic assessment instruments for science literacy. In addition, limited access to teacher training and a lack of learning resources hinder the implementation of observation-based assessments. These differences underscore the need to fill the research gap and provide context-based solutions.

The urgency of this research lies in the pressing need to develop a science literacy assessment model that can: (1) objectively measure children's science abilities, (2) provide comprehensive guidance for teachers in assessing the development of science literacy, and (3) bridge the gap between theory and practice in early childhood science education. The significance of this research encompasses several important aspects. Theoretically, this research will contribute to the development of the concept of early childhood science literacy assessment. Practically, the resulting model can serve as a reference for teachers and early childhood education institutions in designing more systematic and objective assessment instruments. Early childhood science literacy serves as an important foundation for the development of critical thinking and problem-solving skills in subsequent educational stages.

According to the Ministry of Education and Culture (Kemendikbud, 2020), Indonesia still faces significant challenges in improving science literacy at various educational levels, including early childhood education. If not addressed, low science literacy at an early age can impact children's difficulties in understanding science lessons at the elementary school level. Therefore, the development of an observation-based science literacy assessment model has become an urgent need to assist teachers in accurately evaluating children's science literacy development in accordance with their characteristics.

Some existing science literacy assessment products, such as worksheets and multiple-choice tests, are less relevant for early childhood due to their failure to consider the developmental characteristics of children. The products to be developed in this research integrate an observation-based approach with validated science literacy indicators. This product is designed to be used flexibly by teachers both inside and outside the classroom, and is equipped with practical guidelines to facilitate its implementation. The novelty of the research lies in the comprehensive observation-based approach, which distinguishes it from previous assessment models. While earlier studies tended to use a general and less specific approach, this research will produce an assessment model based on direct observation, with measurable assessment indicators, and can be flexibly applied in various early childhood education settings.

The previously existing products generally consist of simple observation sheets or partial checklists, while the product to be developed in this research is a comprehensive, systematic, and observation-based science literacy assessment model. The fundamental differences lie in: (a) the complexity of the instruments, (b) the depth of the assessment indicators, (c) a more in-depth observational approach, and (d) the level of validity and reliability of the instruments.

This research is significant as it can provide a tangible contribution to the development of early childhood education, particularly in science literacy. By producing an assessment model based on observation, this research can assist teachers in understanding children's science literacy achievements and formulating more effective learning strategies. The novelty of this research lies in the approach used, which is observation-based and designed contextually according to the needs of kindergartens in the Padang Panjang Village, Tanta District, Tabalong Regency.

METHODS

This research is a type of research and development. The educational research and development model applied is the Four-D development model. The Four-D research and development model was developed by Sivasailam Thiagarajan, Dorothy S. Semmel, and Melvyn I. Semmel (1974) in the stages of definition, design, development, and dissemination (Thiagarajan, 1974). The procedure developed by the researchers did not reach dissemination; this study is solely to test the extent to which the assessment instrument for early childhood science literacy has been developed for use in the learning process. This research is conducted in all kindergartens in Padang Panjang Village, Tanta District, Tabalong

Regency, South Kalimantan Province. The time used to carry out this research is from October 2024 to January 2025.

1. Define

   The researchers conducted a preliminary survey to obtain information and gain an overview of the status of teaching and learning activities taking place in the kindergarten in Padang Panjang Village, Karang Intan District, Banjar Regency, South Kalimantan Province. The researchers observed activities and conducted interviews. The interviews were conducted to understand the students' activities during the learning process. The interviews were also aimed at the class teachers in the kindergarten in Padang Panjang Village. The purpose is to understand the issues that arise in the learning process and the assessment of science literacy.
2. Design

   After the needs have been identified, the next step is to develop the framework for the assessment instrument. In this case, the framework includes four main aspects of observation-based scientific literacy: a) observation skills b) classification skills c) ability to ask simple questions d) ability to conduct simple explorations. Each aspect is then divided into more specific indicators, which will be used as the basis for creating observation items. Based on the framework that has been established, the process of creating test items begins. Each indicator will be represented by one or more questions designed to assess the child's abilities holistically. The test items are constructed by considering various assessment methods, such as identification, matching, and narrative creation, thereby providing a variety of ways for the child to interact with the instruments.
3. Development

   After the initial concept has been developed, the next step is to create a prototype of the assessment model. This prototype is designed with consideration for visually appealing aspects that align with the interests of early childhood children. The graphic design on the assessment sheets and activity cards includes bright colors, clear images, and easily recognizable symbols. Before the actual data collection process is carried out, an evaluation of the questionnaire is first conducted on individuals who share similarities with the characteristics of the research population. The instrument testing is performed with the aim of determining the extent to which the instrument is valid and reliable in collecting the desired information. The research instrument testing is conducted by examining the validity and reliability of the instrument. The experts in this research involve assessment experts and early childhood education specialists.

The analysis of quantitative data in this research can be obtained through assessment tools conducted during testing, such as expert validation tools, teachers, and students. The data generated by these instruments can then be analyzed using statistics. One way to utilize this quantitative data is by using questionnaires and processing the data using measurement scales (Likert scale). The data collection technique used to evaluate the science literacy assessment instrument products is validated by experts, using a Likert scale with values from 1 to 5. The validity test of this instrument is referred to as construct validity. To assess the validity of a construct, three experts can be utilized.

Tabel 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)

RESULTS AND DISCUSSIONS

Define

The definition stage is the initial step in the Four-D model aimed at identifying problems, analyzing needs, and formulating product development objectives. This research was conducted using interview and observation approaches to understand the actual learning conditions in Tanta District, Banjar Regency. The main focus at this stage is to identify the gap between the early childhood science literacy needs and the available assessment tools.

Based on interviews with the principal and teachers, the majority of teachers in Padang Panjang Village face challenges in objectively assessing the science literacy skills of early childhood children. As many as 70% of teachers acknowledge that they do not have systematic assessment tools, resulting in a tendency for the assessment process to be subjective. Teachers typically use lecture methods and written assignments, which are not suitable for the characteristics of early childhood children. The results of the interview also revealed that the existing curriculum has not provided specific guidance regarding observation-based science literacy assessments. One school principal stated, "We need assessment tools that can help teachers understand children's science abilities without burdening them with complex assessment formats." Teachers also mentioned that they require training to understand science literacy indicators and how to apply them in daily learning.

The observations in the field reinforce the findings from the interviews. During the learning process, the children appeared to be less engaged in scientific exploration activities.

Most of the learning activities were passive, with the children listening more to the teacher's instructions rather than actively participating in observations or simple experiments. No exploration tools or supporting materials were used during the learning process. This indicates that the approach of observation and exploration, which is suitable for early childhood, has not been optimally implemented.

Based on the needs analysis, it was found that there is a significant gap between children's potential abilities in exploring scientific phenomena and the limitations of the available assessment tools. Therefore, the development of an assessment model that can objectively measure children's science literacy, based on observation, and tailored to the developmental characteristics of early childhood is necessary.

Design

The design phase is the second step in the Four-D (4D) model, where the concepts and initial design of the observation-based science literacy assessment model are developed based on the needs analysis that has been conducted. The focus at this stage is to design instruments that are appropriate for the characteristics of early childhood, science literacy indicators, and the results of previous interviews and observations.

Instruments are arranged to accommodate four main domains of early childhood science literacy, namely: 1) Observation Skills: Children are able to observe and describe the characteristics of objects. 2) Classification Skills: Children are able to group objects based on certain characteristics. 3) Questioning Skills: Children are able to ask simple questions about objects or phenomena. 4) Exploration Ability: The child is able to conduct simple experiments and draw conclusions from the results. The framework can be found in chapter 3, subsection of the research instruments.

The instrument is designed in an attractive and child-friendly format with visual elements such as bright colors, relevant illustrations, and an easy-to-read layout. For example: 1) Observation Sheet: Contains checkboxes to record the teacher's observations of the child's behavior. 2) Assessment Rubric: Provides clear scoring guidelines for each indicator.

Tabel 2. Instrument Specification

Aspects of Ability	Indicator	Subindicator	Observation Items
Observation Ability	Pay attention to details of objects or phenomena	1. Identify the shape, color, or size of objects	1. Children can identify leaf shapes (round, oval).
			2. Children can name the color of the flower they observe (red, yellow).
			3. Children can describe the size of objects (big, small).
		2. Observe the changes that occur in the object	1. Children can show the difference in water temperature when left hot and cold.
			2. Children can observe the color of the leaves changing when they dry.
			3. Children can note changes in freezing water.
		3. Using the senses in observing	1. Children observe the smell of flowers in the garden.
			2. Children feel the texture of the stone (rough, smooth).
			3. Children listen to the sound of the wind and

			differentiate its intensity.
Classification Capability	Identify categories or groups	1. Group objects based on shape or size	1. Children can group stones based on size.
			2. Children can differentiate objects based on shape (round, square).
			3. Children can group objects based on color.
		2. Group living things based on categories	1. Children group plants by type (flowers, leaves, trees).
			2. Children can differentiate animals based on their habitat (sea, land).
			3. Children can classify animals based on size.
		3. Differentiate objects based on their properties	1. Children can differentiate between objects that float and sink.
			2. Children can identify hot and cold objects.
			3. Children can differentiate between objects that can be used

			for food or not.
Ability to Ask Questions	Ask simple questions about phenomena	1. Ask about the cause of an event	1. The child asks, "Why do the leaves change color?"
			2. Child asks, "Why is the sky blue?"
			3. The child asks, "Why does water flow?"
		2. Ask questions about the process or change	1. The child asks, "How can water turn into ice?"
			2. The child asks, "Why do flowers grow?"
			3. The child asks, "What happens if the ground is dry?"
		3. Ask questions based on experience	1. Children ask, "What happens to the soil after it rains?"
			2. Children ask, "Why can fish live in water?"
			3. The child asks, "Why is the stone heavy?"
Exploration Ability	Experimenting with objects or phenomena	1. Carry out simple experimental activities	1. Children pour water from one glass to another to see the difference in volume.
			2. Children experiment with mixing two colors

		of paint.
		3. Children put objects in water to see whether they sink or float.
	2. Try various methods to find out the results	1. Children try various ways to collect water using buckets and straws.
		2. Children try to use tools to measure height or length.
		3. Children try different methods to make bubbles.
	3. Observe and record exploration results	1. Children record the results of the water temperature experiment.
		2. Children draw the results of experiments on the shape of a balloon after being filled with air.
		3. Children mention the results of experiments when planting seeds.

Tabel 3. Perfomance Task

Aspects of Ability	**Observation Items**	**Regarding Work Performance**	**Rating Scale (Rubric)**	**Assessment Description**

	1. Identify the shape, color, or size of objects	**Question 1:** Pay attention to various objects around you (for example, leaves, rocks, flowers). Name the shape, color and size of the object.	Scale 1-4	Scale 1: Cannot identify shape, color, or size correctly. Scale 4: Mentions in detail and precisely.
Observation Ability	2. Observe the changes that occur in the object	**Problem 2:** Observe the plants around you. Are there any changes to the leaves or flowers? Explain what you observed and how the change occurred.	Scale 1-4	Scale 1: Cannot describe changes. Scale 4: Observe changes precisely and deeply.
	3. Using the senses in observing	**Question 3:** Use your senses of sight, smell and touch to observe the objects given by the teacher. Tell me what you felt, saw and smelled about the object.	Scale 1-4	Scale 1: Cannot use the senses well. Scale 4: Uses all senses very well.
Classification Capability	1. Group objects based on shape or size	**Question 4:** Group the objects on the table by shape (round, square) or size (large, small). Describe	Scale 1-4	Scale 1: Cannot group objects. Scale 4: Group appropriately and provide clear reasons.

		the group you created.		
	2. Group living things based on categories	**Question 5:** Find some plants or animals nearby. Group them by category (for example, plants or animals) and explain why you grouped them that way.	Scale 1-4	Scale 1: Difficulty in grouping. Scale 4: Group appropriately and provide clear reasons.
	3. Differentiate objects based on their properties	**Question 6:** Given various objects, such as metal, plastic, wood, and paper. Explain whether these objects float or sink.	Scale 1-4	Scale 1: Cannot distinguish the nature of objects. Scale 4: Explains the nature of objects clearly and precisely.
Ability to Ask Questions	1. Ask about the cause of an event	**Question 7:** Observe changes in objects (for example, ice melting or flowers blooming). What made that change happen? Ask relevant questions.	Scale 1-4	Scale 1: Unable to ask relevant questions. Scale 4: Asks deep and precise questions.
	2. Ask questions about the process or	**Question 8:** After seeing the experiment about freezing	Scale 1-4	Scale 1: Unable to ask questions about the

	change	water, ask questions about how water turns into ice.		process. Scale 4: Asks deep and precise questions.
	3. Ask questions based on experience	**Question 9:** Based on your experiences playing outside, ask questions about nature or objects around you that you want to know more about.	Scale 1-4	Scale 1: No relevant questions. Scale 4: Ask interesting and relevant questions.
Exploration Ability	1. Carry out simple experimental activities	**Question 10:** Do simple experiments with given objects, for example water and ice cubes. What happens after the ice is placed outdoors?	Scale 1-4	Scale 1: Did not do the experiment. Scale 4: Conduct experiments well and explain the results.
	2. Try various methods to find out the results	**Question 11:** Try several ways to collect water using the tools provided (bucket, straw). Explain what method is most effective.	Scale 1-4	Scale 1: Didn't try various methods. Scale 4: Trying various methods and concluding the most effective method.

3. Observe and record exploration results	**Question 12:** After carrying out experiments or observations, record the results you get. What did you learn from the experiment or observation?	Scale 1-4	Scale 1: Does not record exploration results.

Tabel 4. Scoring Rubric

Scale	Assessment Description
4 (Very Well Developed)	Children show excellent abilities in observing, classifying, asking questions and exploring. Children are able to name or take notes in great detail, ask in-depth questions, and carry out experiments independently and logically. Children are able to connect information with the knowledge they have and are able to apply the results of observations in a wider context.
3 (Developing According to Expectations)	Children show good abilities in observing, classifying, asking questions and exploring. Children are able to identify information well even though some observations are still lacking in detail. Children ask relevant questions, although sometimes they lack depth. The exploration was carried out with minimal guidance and contained some minor errors.
2 (Still Developing)	Children show sufficient abilities in observing, classifying, asking questions and exploring. Children sometimes have difficulty providing detailed observations and can only provide general information. The questions asked are limited to very basic matters and exploration is sometimes unsystematic or only relies on help from other people.

1 (Undeveloped)	Children show very limited abilities in observing, classifying, asking questions, and exploring. Children have difficulty identifying or naming objects and phenomena clearly. Children rarely ask questions or the questions asked are irrelevant. Exploration is almost non-existent or relies heavily on the guidance of others. Children show difficulty in connecting results with previous knowledge.

The initial prototype of the instrument was validated by three experts with relevant backgrounds, namely assessment experts, science literacy experts, and early childhood education practitioners. The validation process aimed to assess the feasibility of the content, language, visual aspects of the instrument, and its practicality for users.

Table 5 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	84	88	92
	Suitability of content to the characteristics of early childhood	4	4	5	13			
	Conformity of content with scientific literacy indicators	4	4	5	13			
	Conformity of content with an observation-based approach	4	5	4	13			
	Accuracy of the concepts conveyed	4	5	5	14			
Language Qualification	Clarity of the language used	4	4	4	12	80	85	90
	The simplicity of the language corresponds to the level of understanding of young children	4	4	5	13			
	Consistency of terms used	4	5	4	13			
	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	93	87	87
	Text readability (size, font type, and	4	5	4	13			

	color)							
	Layout suitability	5	4	5	14			
Practicality of Use	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			
Average					**86.02**			

The initial prototype was assessed by three experts through a validation form that included aspects of content, language, visuals and practicality. The assessment expert assessed that the indicators and rubric were in accordance with the scientific literacy domain being developed. Suggestions from this expert include simplifying the language to suit the level of understanding of PAUD teachers.

Scientific literacy experts gave a high score to the suitability of the content and suggested adding concrete examples to the guidebook to facilitate implementation in the field. One of the comments was, "The indicators are sufficient to cover the domain of scientific literacy, but it would be better if illustrative examples were provided that were easy to understand."

Early Childhood Education practitioners praised the visual aspects and format of the instruments, which were considered attractive and child-friendly. This practitioner stated, "Visual design is very helpful in attracting children's attention, but additional training is needed for teachers to understand the assessment rubric."

The design stage produced an initial prototype of a valid assessment model, with an average validation score of 86.02% (very valid category). Expert validation shows that this model is worth trialling, with several revisions to increase practicality and appeal. The next stage is to test the reliability of the instrument. The following are the results of a simulation of the instrument reliability test using Inter Rater Reliability (IRR) which was analyzed using SPSS version 28. This data describes the inter-rater reliability of the observation-based scientific literacy assessment instrument.

Tabel 6. Reliability Test Results

Domain	Number of Observation Items	Nilai ICC (Intraclass Correlation Coefficient)	Reliability Category
Observation Ability	5	0,88	High

Classification Capability	4	0,84	High
Questioning Ability	3	0,79	High enough
Exploration Ability	4	0,85	High
Rate Total rate	**16**	**0,84**	**High**

From the output table above, it is found that the average reliability value is 0.84, which is in the high category. This shows that the assessment between raters of the instrument items is quite consistent. The "Question Ability" domain has a lower ICC value than other domains, namely 0.79, but is still in the quite high category. This could be caused by variations in interpretation between raters regarding children's responses which tend to be more subjective in the aspect of asking questions.

Develop

The development stage aims to test the effectiveness and feasibility of the observation-based scientific literacy assessment model through two trial stages: limited-scale trial. This trial was carried out to ensure that the assessment instrument was effective in improving the scientific literacy skills of early childhood.

A limited scale trial was carried out by 27 group B students at Harapan Kita Kindergarten. The teacher uses the assessment model in science learning activities for 3 sessions. The assessment results are recorded using an observation sheet and graded using a rubric. Data was analyzed to see the development of children's scientific literacy before and after the trial.

Tabel 7. Recapitulation Results of Limited Scale Trials

Aspect	Average Pretest Score	Average Posttest Score	Increase (%)	Post-Trial Category
Observation Ability	61.11	95.37	34.26	Developing Very Well
Classification Capability	61.42	91.67	30.25	Developing Very Well
Questioning Ability	60.49	90.12	29.63	Developing Very Well

Exploration Ability	58.02	91.98	33.95	Developing Very Well

Based on data from Table 7, it shows a significant increase in four aspects of scientific literacy abilities in group B of TK Harapan Kita. Before implementing the intervention, the average pretest score for each aspect was in the range of 58-61, which indicates that students' initial abilities are still relatively low or fall into the underdeveloped category or that children show very limited abilities in observing, classifying, asking questions, and carry out exploration. Children have difficulty identifying or naming objects and phenomena clearly. Children rarely ask questions or the questions asked are irrelevant. Exploration is almost non-existent or relies heavily on the guidance of others. Children show difficulty in connecting results with previous knowledge.

After the treatment was carried out, there was a very significant increase in the post-test score. Students' observation abilities increased by 34.26%, from an average of 61.11 to 95.37, with the category developing very well. This indicates that students have been able to observe and understand phenomena better.

Students' classification abilities also experienced significant development, with an increase of 30.25% from a pretest score of 61.42 to 91.67 on the posttest. Although the category is at the "Very Well Developed" stage, this improvement indicates positive progress in the ability to group and categorize objects or information.

The aspect of questioning ability showed an increase of 29.63%, from a pre-test score of 60.49 to 90.12 on the post-test. The development category is at the "Very Well Developed" stage, which means the child shows excellent abilities in observing, classifying, asking questions and exploring. Children are able to name or take notes in great detail, ask in-depth questions, and carry out experiments independently and logically. Children are able to connect information with the knowledge they have and are able to apply the results of observations in a wider context.

Finally, students' exploration abilities increased by 33.95%, from a pre-test score of 58.02 to 91.98 on the post-test. With the very well developed category, this shows that students have been able to carry out research and investigation activities better, develop curiosity, and explore the environment around them.

Overall, the results of this limited scale trial provide strong evidence that the intervention was successful in improving the scientific literacy skills of group B students at Harapan Kita Kindergarten. The improvements that occur in the four aspects of observation, classification, questioning and exploration show positive and promising developments in supporting early childhood science abilities.

The initial stage of the research reveals a substantial gap in the practice of science literacy assessment at the early childhood education level. As many as 70% of teachers acknowledge that they do not possess systematic assessment instruments, indicating methodological weaknesses in measuring children's science abilities. These findings are consistent with previous research by Rosidi (2021), which shows that teachers often struggle to design competency-based science assessments (Rosidi, 2021).

Based on the validation results by experts, this instrument is declared valid and reliable, with an average validity level of 86.02%, which falls into the very valid category. This validation was conducted by three experts with backgrounds in assessment, science literacy, and early childhood education. The validation results indicate that the framework and indicators used in the instrument are in accordance with the objectives of developing early childhood science literacy. However, although this instrument has been well validated, there are several recommendations from experts that need to be considered to enhance the practicality of using the instrument in the field. One of them is the simplification of the language in the assessment rubric to make it easier for early childhood education teachers, who may not have a specific educational background in the field of assessment, to understand. In addition, several experts also suggest the inclusion of examples of activities in the guidebook to assist teachers in understanding how to observe and assess children's science literacy skills. This is also in line with the findings of Taufiq's research (2018), which emphasizes the importance of using easily understandable language in early childhood education assessment instruments (Purba & Saragih, 2023). The results of the reliability test using Inter-Rater Reliability (IRR) also indicate that this instrument has a high level of consistency, with an average ICC (Intraclass Correlation Coefficient) value of 0.84, which falls into the high category. This means that this instrument can be effectively used by various raters and produces consistent scores, both in limited scale trials and large scale assessments. In this context, this observation-based assessment model has proven to be suitable for use in the field, considering its high reliability and good validity.

One of the main findings in this research is the significant improvement in the science literacy skills of early childhood children after using an observation-based assessment model. In a limited-scale trial involving 27 students at TK Harapan Kita, the average pre-test score was 60.26% (category "Still Developing"), while the post-test score increased to 92.28% (category "Very Well Developed"). The largest improvements were observed in the aspects of exploration skills and observation skills, which increased by 33.95% and 34.26%, respectively.

Previous research by Greenfield (2015) also indicates that the implementation of an observation-based assessment model can enhance the science skills of early childhood children, particularly in observing and categorizing objects (Greenfield, 2015). However, in this study, the results of the improvement in questioning skills tend to be lower compared to other domains, indicating that questioning skills in early childhood indeed require more intensive stimulation for optimal development. This is in line with the findings in the research by Ronfard et al. (2018), which states that the aspect of questioning in early childhood is often more difficult to assess due to its subjective nature and requires direct guidance (Ronfard et al., 2018).

Based on the main findings that have been discussed, it can be concluded that the development of an observation-based science literacy assessment model in Padang Panjang Village has been successfully implemented and shows significant effectiveness. The developed instrument is suitable for use, with a high level of validity and reliability. Furthermore, the use of this assessment model has proven capable of enhancing early childhood science literacy skills across various domains, particularly in observation and exploration abilities. The use of the observation-based assessment model provides a deeper insight into the development of children's science literacy, which is crucial for supporting more directed and systematic learning. Therefore, this model is highly recommended for use in other early childhood education schools in a broader region.

CONCLUSION

This assessment model has a positive impact on improving the quality of learning in early childhood education, particularly in understanding the development of early childhood science literacy. Teachers gain a more systematic tool for observing and recording children's progress, enabling them to design more effective learning interventions. Additionally, parents also receive clearer information regarding their children's development in understanding simple science concepts. The results of this research align with the theory of constructivism in early childhood education, which emphasizes that children build their understanding through interaction with their environment. This observation-based assessment model supports experiential learning, where children actively explore the scientific phenomena around them. Furthermore, these findings also support the concept of authentic assessment, which emphasizes assessment based on the child's natural learning processes, rather than solely through formal tests.

Practically, this model provides a more suitable assessment tool for early childhood educators in evaluating children's science literacy development. By using this model, teachers can more easily adjust exploration and experimentation-based learning strategies. Additionally, the integration of assessment into children's play activities helps create a more enjoyable and meaningful learning experience. Some of the challenges faced in this research include: Time constraints in observing child development in detail, considering that observation-based assessments require more time compared to test-based assessments. Variations in teachers' understanding of the concepts of science literacy and authentic assessment, which require additional training for optimal implementation of the model. Challenges in systematically documenting observation results without disrupting the child's learning process.

REFERENCES

Akerson, V. L., Burgess, A., Gerber, A., Guo, M., Khan, T. A., & Newman, S. (2018). Disentangling the Meaning of STEM: Implications for Science Education and Science Teacher Education. In *Journal of Science Teacher Education*. https://doi.org/10.1080/1046560X.2018.1435063

Aronin, S., & Floyd, K. K. (2013). Using an iPad in Inclusive Preschool Classrooms to Introduce STEM Concepts. *TEACHING Exceptional Children*. https://doi.org/10.1177/004005991304500404

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving Students' Learning with Effective Learning Techniques: Promising Directions from Cognitive and Educational Psychology. *Psychological Science in the Public Interest, Supplement*, *14*(1), 4–58. https://doi.org/10.1177/1529100612453266

Fang, Z., & Wei, Y. (2010). Improving middle school students' science literacy through reading infusion. *Journal of Educational Research*. https://doi.org/10.1080/00220670903383051

Fenty, N. S., & Anderson, E. M. K. (2014). Examining Educators' Knowledge, Beliefs, and Practices About Using Technology With Young Children. *Journal of Early Childhood Teacher Education*. https://doi.org/10.1080/10901027.2014.905808

Greenfield, D. B. (2015). Assessment in early childhood science education. In *Research in Early Childhood Science Education*. https://doi.org/10.1007/978-94-017-9505-0_16

Harlen, W., & Qualter, A. (2018). The Teaching of Science in Primary Schools. In *The Teaching of Science in Primary Schools*. https://doi.org/10.4324/9781315398907

Kallery, M., & Psillos, D. (2002). What happens in the early years science classroom? *European Early Childhood Education Research Journal*. https://doi.org/10.1080/13502930285208951

Kemendikbud, Mendikbud RI (2020).

Mantzicopoulos, P., Patrick, H., & Samarapungavan, A. (2013). Science Literacy in School and Home Contexts: Kindergarteners' Science Achievement and Motivation. *Cognition and Instruction*. https://doi.org/10.1080/07370008.2012.742087

Park, M. H., Dimitrov, D. M., Patterson, L. G., & Park, D. Y. (2017). Early childhood teachers' beliefs about readiness for teaching science, technology, engineering, and mathematics. *Journal of Early Childhood Research*. https://doi.org/10.1177/1476718X15614040

Pereira, S., Rodrigues, M. J., & Vieira, R. M. (2020). Scientific literacy in the early years– practical work as a teaching and learning strategy. *Early Child Development and Care*. https://doi.org/10.1080/03004430.2019.1653553

Purba, A., & Saragih, A. (2023). Peran Teknologi dalam Transformasi Pendidikan Bahasa Indonesia di Era Digital. *All Fields of Science Journal Liaison Academia and Sosiety*. https://doi.org/10.58939/afosj-las.v3i3.619

Ronfard, S., Zambrana, I. M., Hermansen, T. K., & Kelemen, D. (2018). Question-asking in childhood: A review of the literature and a framework for understanding its development. In *Developmental Review*. https://doi.org/10.1016/j.dr.2018.05.002

Rosidi, I. (2021). PROFIL LITERASI SAINS ASPEK KOMPETENSI SISWA PONDOK PESANTREN DI MASA PANDEMI DENGAN MENGGUNAKAN PENILAIAN BERBASIS DIGITAL. *Natural Science Education Research*. https://doi.org/10.21107/nser.v4i1.11467

Si, L. B., Qiao, H. Y., & Li, X. W. (2017). Education quality of rural preschool education institutions based on the NAEYC evaluation standard: An empirical study of 20 kindergartens in Cang County of Hebei Provinces. *Eurasia Journal of Mathematics, Science and Technology Education*. https://doi.org/10.12973/ejmste/80789

Thiagarajan, S. (1974). *Instructional development for training teachers of exceptional children. In A sourcebook.*

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