Jurnal MADINASIKA
Homepage: https://ejournal.unma.ac.id/index.php/madinasika
Vol. 7 No. 2, Bulan April 2026, halaman: 390~404
E-ISSN: 2716-0343, P-ISSN: 2715-8233
http://dx.doi.org/10.31949/madinasika.v7i2.17872

EVALUATION OF THE IMPLEMENTATION OF LEARNING OUTCOME TESTS IN PAI-BP LEARNING IN CLASS VI OF SDN 5 MENTENG PALANGKA RAYA

Ihsan Auliya Arrasyid [1*], Abdul Qodir [2], Aghnaita Aghnaita [3]
[1,2,3] Pendidikan Agama Islam, Universitas Islam Negeri Palangka Raya, Palangka Raya, Kalimantan Tengah, Indonesia
[1*]Email penulis koresponden: ihsanaa2211110123@uin-palangkaraya.ac.id

Riwayat Artikel

Submited:
03-04-2026
Accepted:
18-04-2026
Published:
18-04-2026

Abstrak

Penelitian ini bertujuan untuk mengevaluasi pelaksanaan tes hasil belajar Pendidikan Agama Islam dan Budi Pekerti (PAI-BP) di kelas VI SDN 5 Menteng Palangka Raya menggunakan model evaluasi CIPP (*Context, Input, Process, Product*) dengan pendekatan kuantitatif. Penelitian ini merupakan penelitian evaluatif dengan subjek siswa kelas VI yang ditentukan secara purposive. Pengumpulan data dilakukan melalui angket skala Likert empat poin, observasi menggunakan lembar ceklis, serta dokumentasi, kemudian dianalisis secara deskriptif kuantitatif. Hasil penelitian menunjukkan bahwa pelaksanaan tes hasil belajar PAI-BP pada aspek context memperoleh nilai rata-rata angket sebesar 3,36 dengan kategori tinggi dan observasi sebesar 3,20 dengan kategori baik. Aspek input memperoleh nilai rata-rata angket 3,37 tinggi dan observasi 3,89 sangat baik. Aspek process menunjukkan nilai rata-rata angket 3,34 tinggi dan observasi 3,60 sangat baik. Aspek product memperoleh nilai rata-rata 3,39 tinggi dan observasi sebesar 2,89 dengan kategori baik. Model CIPP efektif memberikan gambaran komprehensif sebagai dasar perbaikan evaluasi pembelajaran.

Kata kunci: Evaluasi Pembelajaran; Model CIPP; PAI-BP; Tes Hasil Belajar.

Abstract

This study aims to evaluate the implementation of Islamic Religious Education and Morality (PAI-BP) learning outcome tests in grade VI of SDN 5 Menteng Palangka Raya using the CIPP (Context, Input, Process, Product) evaluation model with a quantitative approach. This study is an evaluative study with grade VI students as subjects, who were selected purposively. Data collection was conducted through a four-point Likert scale questionnaire, observation using a checklist, and documentation, then analyzed descriptively and quantitatively. The results showed that the implementation of the PAI-BP learning outcome test in the context aspect obtained an average questionnaire score of 3.36 in the high category and an observation score of 3.20 in the good category. The input aspect obtained an average questionnaire score of 3.37, categorized as high, and an observation score of 3.89, categorized as very good. The process aspect showed an average questionnaire score of 3.34, categorized as high, and an observation score of 3.60, categorized as very good. The product aspect obtained an average score of 3.39, categorized as high, and an observation score of 2.89, categorized as good. The CIPP model effectively provides a comprehensive overview as a basis for improving learning evaluation.

Keywords: *Learning Evaluation; CIPP Model; PAI-BP; Learning Outcome Tests.*

Jurnal ***MADINASIKA*** *diterbitkan oleh Fakultas Pascasarjana, Program Studi Magister Manajemen Pendidikan Islam, Universitas Majalengka*

INTRODUCTION

Given today's ever-evolving social and educational landscape, the education sector faces growing challenges in preparing students not only to meet academic demands but also to fulfill moral and ethical expectations. Social changes, shifting values, and the increasing complexity of students' environments require the education system to place greater emphasis on character development alongside cognitive achievement. One form of education that should be provided to strengthen moral understanding, attitudes, and character particularly among Muslim students is Islamic religious education. Religious education is considered to play a vital role in fostering the morals and culture of a dignified nation (Khasanah et al., 2024). Islamic religious education aims not only to instill religious teachings, but also to shape students' personalities and character in accordance with Islamic values (Sapitri & Maryati, 2022). Strengthening religious education is a strategic necessity in developing a generation that is not only intellectually intelligent, but also spiritually and socially superior.

However, the success of PAI-BP in addressing these challenges depends not only on the delivery of instruction but also on how effectively learning outcomes are assessed. Assessment, particularly through learning outcome exams, is essential for determining whether students have truly understood and internalized the values and knowledge taught. To ensure that these values are truly internalized in students, an effective mechanism is needed in the learning process. This mechanism can be in the form of learning outcome tests, which serve to measure the achievement of learning materials and understand the extent to which students have mastered PAI-BP materials (Fitrianti, 2018). Learning outcome tests serve as periodic assessment instruments that measure students' understanding during the learning process (Hasmawati & Muktamar, 2023). Teachers can assess students' level of understanding of the religious values learned through these learning outcome tests and provide the necessary feedback for improving and strengthening the quality of learning.

The results of research by show that the implementation of assessment in PAI-BP subjects at SMP Negeri 3 Hulu Kuantan went well overall. Of the eight indicators implemented in the assessment per unit of learning material, only one indicator was not fully in accordance with the theory, while of the seven assessment indicators at the end of the learning period, there were two indicators that were not optimally implemented. This is in line with the findings of Wildana et al., (2024) who stated that the implementation of learning assessment in PAI learning at SD Negeri 02 Kampung Baru applied a contextual and participatory approach involving initial assessment, assessment during the learning process, and final evaluation. The implementation of this assessment proved to be able to support the development of students' academic, social, and character abilities. However, there are still several obstacles, such as limited learning time and classroom conditions that affect the optimization of learning outcome tests. Taqiyuddin et al., (2024) stated that PAI in schools has the main objective of shaping students who are religious and obedient in practicing their religious teachings, with assessments in learning serving as a tool to assess and measure the abilities of each student.

Although various studies show that learning outcome tests can contribute to improving the quality of learning, their implementation still needs serious attention because they are closely related to the ability of learning outcome tests to provide constructive feedback for students and teachers (Munaroh, 2024). Ideally, learning outcome tests can help teachers adjust their learning strategies according to students' needs. However, there are still obstacles in their implementation, such as time constraints, lack of follow-up on the test results, and so on (Oktiviana et al., 2024). The types of learning outcome tests that are usually applied include multiple choice questions, short answer questions, essays, and others. These forms are used to explore various aspects of student understanding, ranging from cognitive, affective, and psychomotor knowledge in the application of religious values in students' daily lives (Rahman & Nasryah, 2019).

The use of learning outcome test results as a basis for improving learning strategies still needs to be reviewed in more depth. Many teachers only use learning outcome tests as a routine assessment tool without using them to comprehensively analyze students' learning difficulties. If learning outcome tests are not analyzed properly, gaps in student understanding cannot be detected effectively (Ratnawulan & Rusdiana, 2015). Learning evaluation is an important

component in a learning or educational process that is carried out to determine the level of achievement or value of students during teaching and learning activities in a period (Tsawab et al., 2019). herefore, the evaluation of the form and implementation of learning outcome tests covers the extent to which the questions provided are able to measure student understanding.

Evaluation of the implementation of learning outcome tests is important to review the extent to which these tests are in line with the curriculum being implemented (Adinda et al., 2021). The curriculum requires the implementation of a more contextual approach in PAI-BP (Akbar et al., 2024). This evaluation serves to identify weaknesses in the teaching materials so that they can be continuously improved (Indriani et al., 2021). Therefore, the learning outcome tests used should not only emphasize memorization but also measure conceptual understanding and students' ability to implement religious values in their daily lives.

Evaluating the implementation of learning outcome tests in PAI-BP learning at SDN 5 Menteng is a very interesting thing to do. This evaluation will provide an overview of how these tests can be optimized to better suit learning objectives. One evaluation model that can be used is the CIPP (Context, Input, Process, Product) model, which allows for a systematic analysis of the implementation of learning outcome tests from various aspects. Context evaluation focuses on identifying the needs underlying the program objectives. Input evaluation assists in strategic decision-making and program design planning. Process evaluation is used to detect weaknesses in program implementation and make improvements during implementation. Product evaluation is aimed at assessing the results or outputs of the program that has been implemented (Wirawan, 2016). The results of this evaluation will also review the extent to which the test can measure student understanding more effectively.

Although previous studies have examined the implementation of PAI-BP learning assessment and evaluation, most of these studies still focus on the general effectiveness of assessment or the description of assessment practices without comprehensively examining the relationship between learning objectives, test design, implementation process, and the use of test results as a basis for systematic learning improvement. revious studies tend to place learning outcome tests as part of general learning evaluation, without specifically evaluating the implementation of PAI-BP learning outcome tests using a comprehensive CIPP evaluation model approach, especially at the elementary school level. The uniqueness of this study lies in the use of the CIPP evaluation model to comprehensively analyze the implementation of PAI-BP learning outcome tests, starting from contextual suitability to the product in the form of learning test results. This study also provides empirical contributions by placing learning outcome tests not only as a tool to measure cognitive achievement but also as a strategic instrument in supporting PAI-BP learning.

Based on the above description, this study aims to evaluate the implementation of learning outcome tests in PAI-BP learning at SDN 5 Menteng Palangka Raya to measure and improve students' understanding. This evaluation employs the CIPP model to comprehensively assess the implementation of learning outcome tests in grade VI, covering aspects of context, input, process, and product. Theoretically, this study contributes to the development of educational evaluation, particularly in strengthening the application of the CIPP model within the context of Islamic Religious Education (PAI-BP). This study enriches the existing literature by providing empirical evidence on how the CIPP model can be utilized not only to evaluate learning programs but also to assess the effectiveness of achievement tests as part of the evaluation system. Practically, this study is expected to provide useful insights for teachers and schools in designing, implementing, and utilizing learning outcome tests more effectively and reflectively. These findings are expected to help teachers improve their assessment practices, particularly in using test results as a basis for instructional improvement and decision-making. Additionally, this study aims to identify more effective strategies for administering achievement tests, thereby making a positive contribution to students' academic development.

RESEARCH METHOD

This study uses a quantitative approach with an evaluative type using the CIPP model developed by Stufflebeam, which aims to evaluate the implementation of PAI-BP learning outcome

tests (Yuniarti et al., 2021). his model is used to evaluate the implementation of learning outcome tests in PAI-BP learning comprehensively from the aspects of context, input, process, to product to describe empirical phenomena objectively through numerical measurements obtained from student responses and field observation results.

The population in this study was all classes at SDN 5 Menteng Palangka Raya. This population was relevant because it was the main object of focus in this study. The researcher determined the sample purposively, namely by deliberately selecting the most relevant class at SDN 5 Menteng Palangka Raya, which was grade VI.

Data were collected from three types of instruments, namely questionnaires, checklist observation sheets, and documentation. Questionnaires were used to obtain quantitative data on respondents' perceptions of the implementation of learning outcome tests. Questionnaires were given to students to assess the implementation of learning outcome tests from the aspects of context, input, process, and product, which were compiled based on the CIPP model indicators.

The questionnaire was a four-point Likert scale, ranging from very low to very high. The questionnaire was developed by identifying indicators based on the CIPP model, compiling statements, and validating them by experts to ensure the instrument's feasibility. The data obtained through the questionnaire was then analyzed using SPSS with quantitative descriptive analysis based on score conversion. The results obtained were then classified into evaluation assessment criteria according to Mardapi (2019). Each component of the evaluation results was categorized into four levels on a table.

Table 1. Evaluation Assessment Criteria

Interval	Criteria
$X \geq \bar{X} + 1{,}5\ SBx$	Very High
$\bar{X} + 1{,}5\ SBx > X \geq \bar{X}$	High
$\bar{X} > X \geq \bar{X} - 1{,}5\ SBx$	Low
$X < \bar{X} - 1{,}5\ SBx$	Very Low

Description: X = *achieved score*
$\bar{X}$ = *overall average score*
SBx = *standard deviation of overall scores*

The use of these criteria provides a more accurate and proportional picture of the quality of the learning outcome test implementation in each aspect of the CIPP model. The observation data was compared with the questionnaire results to see the consistency of the findings, thereby producing a stronger evaluative interpretation. The entire analysis process was carried out systematically, allowing the research to be replicated and scientifically accountable.

Meanwhile, the researchers used observation sheets in the form of a 1 to 4 scale checklist, which was designed to directly observe the process of administering learning outcome tests in the classroom. The 1–4 observation scale describes the level of implementation, ranging from not good to very good. This observation was conducted to check the suitability of the test implementation with the ideal procedure, including the readiness of the equipment, student conditions, and the flow of the test implementation directly. Each score obtained was analyzed by calculating the total score, calculating the average score, and converting the score based on predetermined criteria. The average obtained is then classified into four categories of implementation level based on the criteria from (Pimentel, 2019).

Table 2. Observation Assessment Criteria

Average Interval	Category
3.28 – 4.00	Very Good
2.52 – 3.27	Good
1.76 – 2.51	Not So Good
1.00 – 1.75	Not Good

The use of these categories provides a more systematic picture of the level of implementation of the learning outcome test process. Documentation is used to complement and strengthen research findings through the search for relevant supporting documents, such as the

questions used and the assessment results. The data analysis in this study uses the mean analysis technique to determine the average score of each evaluation aspect.

RESULT AND DISCUSSION

The term evaluation is often known in English as "evaluation" from the root word "value," which means worth or price. This term was then absorbed into Indonesian as "evaluasi." (Pamulatsih & Zulfitria, 2024). According to (2003 in Wirawan, 2016, p. 10) states that:

> "*Evaluation is the process of delineating, obtaining, and providing descriptive and judgmental information about the worth and merit of some object's goals, design, implementation, and impact in or-der to guide decision making, serve needs for accountability, and promote understanding of the involved phenomena.*"

This definition forms the basis for the development of various evaluation models, one of which is the CIPP model, which is most widely used in educational research. The CIPP evaluation model was first developed by Daniel Stufflebeam and includes four main components, namely context evaluation, input evaluation, process evaluation, and product evaluation. Context evaluation focuses on identifying the initial needs that form the basis for setting the objectives of a program. Input evaluation is used as a basis for strategic decision-making and program design planning. Process evaluation is used to detect weaknesses in program implementation and make improvements during implementation. Product evaluation is aimed at assessing the results or outputs of the program that has been implemented (Wirawan, 2016, p. 95).

One of the main instruments in the context of learning evaluation used to obtain information about student learning achievements is tests. The term "test" comes from the Latin (French) "testum," which literally means a container made of clay used to separate precious metals from other elements, such as stones, soil, or sand (Inanna et al., 2021). A test can also be understood as a tool or procedure used in measurement activities containing a collection of questions, statements, or a series of specific tasks that must be completed or solved by students in order to assess various aspects of student behavior (Ropii & Fahrurrozi, 2017). According to Indrakusuma, learning outcome tests are a series of questions or tasks that are systematically designed to determine the level of students' ability to master the material that has been learned (Yuniar et al., 2015).

Based on the explanation of evaluation theory and the previous discussion, learning evaluation cannot be understood from only one aspect of assessment. Learning outcome tests do provide an overview of student achievement, but they do not fully show the success of the learning process. Therefore, evaluation needs to be viewed more comprehensively in the implementation of learning outcome tests by considering the context, input, process, and learning outcomes.

This study obtained data through questionnaires, observation, and documentation. The data obtained from questionnaires and observation were analyzed descriptively and quantitatively by calculating the average value for each indicator, which was then summarized for each aspect to maintain the consistency of the assessment scale and provide a proportional description of each aspect in the CIPP evaluation model. The average scores were then classified based on assessment categories with predetermined interval ranges. Meanwhile, the documentation data was used as supporting data to reinforce and confirm the results obtained from the questionnaires and observations. The research results were presented sequentially based on the CIPP aspects so that the implementation of learning evaluation could be understood systematically and comprehensively.

a. Evaluation of the Context Aspect

The context aspect is the initial aspect evaluated to examine the objectives, needs, readiness, and suitability of the initial conditions in the implementation of learning outcome tests. Context evaluation is conducted to assist in the planning process and program direction by identifying unmet needs, formulating appropriate objectives, and understanding the characteristics of the program environment (Rahman & Nasryah, 2019). According to Stufflebeam, context evaluation is a basic form of evaluation to provide a

foundation for setting educational goals. This evaluation is aimed at identifying and analyzing various conditions that underlie educational programs, so as to be able to distinguish between problems that arise and the real needs that should be met in an educational program. Context evaluation also encourages educational goals that are relevant to the expected needs (Khaerudin, 2022). The results of observations on table 3.

Table 3. Observation Result for the Context Aspect

Observation	**Average**
Observation 1	3,40
Observation 2	3,20
Observation 3	3,00
Overall Average	**3,20**

Based on the results of observations on the context aspect evaluation, an average score of 3.20 was obtained. This score is in the range of 2.52–3.27, which is classified as good. Most of the indicators in the context aspect have been met, although there are still certain indicators whose scores are not yet optimal, requiring further attention so that the evaluation can be carried out more optimally. The results of the questionnaire given to respondents on the context aspect yielded the following scores table 4.

Table 4. Context Aspect Questionnaire Result

Descriptive Statistics

	N	Mean	Std. Deviation
Mean_Context	32	3.36	.34
Valid N (listwise)	32		

The questionnaire results obtained from respondents on the context aspect had an average score of 3.36 with a standard deviation of 0.34. To determine the category of context evaluation results, interval criteria were used based on the previous evaluation assessment criteria, namely by calculating the upper and lower limits of the interval using the following formula.

1. $\bar{X} + 1{,}5\ SBx$
   $3{,}36+(1{,}5\times 0{,}34)=3{,}36+0{,}51=3{,}87$
2. $\bar{X} - 1{,}5\ SBx$
   $3{,}36-0{,}51=2{,}85$

Then, the results of the upper and lower limits of the interval were classified into table.

Table 5. Evaluation Criteria for Context Aspect

Interval	**Criteria**
$X \geq 3{,}87$	Very High
$3{,}87 > X \geq 3{,}36$	High
$3{,}36 > X \geq 2{,}85$	Low
$X < 2{,}85$	Very Low

The average questionnaire score obtained from the context aspect was 3.36, which is in the range of 3.87 > X ≥ 3.36, thus falling into the high category. This finding shows that, in general, the context of the learning outcome test in PAI-BP learning at SDN 5 Menteng Palangka Raya has been positively assessed by students. This indicates that the purpose of the learning outcome test, the suitability of the material with the competencies, and the relevance of the learning outcome test requirements have been understood and felt to be beneficial by students. On the indicator of clarity of evaluation objectives, most respondents gave positive assessments, indicating that students understand that learning outcome tests aim to measure the achievement of competencies that have been learned. This finding

supports the opinion of Laila et al., (2024) that effective learning outcome tests must begin with clear objectives so that the test implementation process truly measures the expected competencies. Good learning outcome tests must be in line with instructional objectives so that the test results can provide an accurate measure of the level of learning achievement attained by students (Ridha et al., 2025). Although the context aspect is relatively good, it still needs improvement, especially in ensuring that students have a common perception of the objectives and expected needs of the learning outcome tests conducted.

b. Evaluation of the Input Aspect

Input evaluation aims to assist the decision-making process in preparing and managing the various resources needed for the program to run optimally (Rahman & Nasryah, 2019). Input evaluation basically aims to assist in determining the right program as the basis for implementing the necessary changes. Various resources, strategies, and alternative actions are analyzed systematically so that the decisions made truly support the improvement and development of the planned program (Kurniawati, 2021).

This input evaluation includes the adequacy of resources, assessing strategies, implementation plans, and the suitability of strategies to student needs. Supporting infrastructure such as classrooms and other facilities are also an important part of input evaluation to support the success of the learning program (Wan Azman, 2025). Through this input evaluation, information is obtained which is used as a basis for decision making to improve or develop learning programs so that they run more effectively (Tsani et al., 2021). The input aspect was analyzed to determine the readiness of resources, the suitability of strategies, the assessment of strategies, and the work plan and schedule set for the implementation of learning outcome tests in PAI-BP learning in grade VI at SDN 5 Menteng Palangka Raya. The results of the observation of the input aspect on Table 6.

Table 6. Observation Result of Input Aspect

Observation	**Average**
Observation 1	4,00
Observation 2	4,00
Observation 3	3,67
Overall Average	**3,89**

The results of observations on the input aspect show an average score of 3.89. This score falls within the range of 3.26–4.00, which is classified as very good. This finding indicates that the input component in the implementation of the learning outcome test in PAI-BP learning in grade VI of SDN 5 Menteng Palangka Raya has been fulfilled very well. The results of the questionnaire given to respondents on the input aspect obtained a value of table 7.

Table 7. Input Aspect Questionnaire Result

Descriptive Statistics

	N	Mean	Std. Deviation
Mean_Input	32	3.37	.30
Valid N (listwise)	32		

Based on the results of the questionnaire given to respondents on the input aspect, an average value of 3.37 with a standard deviation of 0.30 was obtained. To determine the input evaluation result category, interval criteria were used based on the previous evaluation assessment criteria, namely by calculating the upper and lower limits of the interval using the following formula.

1. $\bar{X} + 1{,}5\ SBx$

   $3{,}37+(1{,}5\times 0{,}30)=3{,}37+0{,}45=3{,}82$

2. $\bar{X} - 1{,}5\ SBx$
   3,37−0,45=2,92

Then, the results of the upper and lower limits of the interval were classified into table 8.

Table 8. Input Aspect Evaluation Assessment Criteria

Interval	Criteria
$X \geq 3{,}82$	Very High
$3{,}82 > X \geq 3{,}37$	High
$3{,}37 > X \geq 2{,}92$	Low
$X < 2{,}92$	Very Low

Based on the input aspect assessment criteria table above, the average questionnaire score obtained from the input aspect of 3.37 was in the range of 3.82 > X ≥ 3.37, which falls under the high criteria. This finding shows that, in general, the input from the implementation of the learning outcome test in PAI-BP learning at SDN 5 Menteng Palangka Raya has been positively assessed by students. This shows that the readiness of resources, the appropriateness of strategies, the assessment of strategies, and the work plans and schedules set out in the implementation of learning outcome tests in PAI-BP learning in grade VI at SDN 5 Menteng Palangka Raya are adequate to support the implementation of learning outcome tests.

The questionnaire results on the indicators of appropriateness and assessment of strategies show that the test implementation strategy is considered to be quite appropriate for the characteristics of Islamic education material. This indicates that teachers have chosen a test format that is relevant to the teaching material. However, the variation in respondents' scores shows that there are still students who feel that this learning outcome test strategy does not fully accommodate differences in learning abilities. Learning outcome tests designed according to competencies will help teachers obtain objective information about students' abilities. (Munandar et al., 2023).

The work plan and test schedule indicators also received relatively high scores, indicating that the timing of the learning outcome tests had been planned quite well. However, the questionnaire results showed that some students still felt that the time allocated for the tests was not entirely proportional. This indicates the need for further evaluation of the time allocation so that the learning outcome tests can be conducted optimally for all students' abilities. (Armedi, 2025).

c. Evaluation of the Process Aspect

Process evaluation serves to directly monitor the implementation of a program so that it is carried out in accordance with the predetermined plans and procedures (Rahman & Nasryah, 2019). Process evaluation in the CIPP model is also an assessment of program implementation, which includes what programs are carried out, who is responsible for the program, when the program is carried out, the objectives of monitoring implementation procedures, supporting decision-making, and documenting the program's progress (Andriani et al., 2021). The purpose of this evaluation is to ensure that the entire process runs optimally in accordance with the plan, while identifying obstacles that arise during implementation so that continuous improvements can be made (Isnaeni et al., 2024). The process aspect evaluation in this study examines the consistency of implementation, stakeholder involvement, program progress records and reports, and the accuracy of the target group in conducting learning outcome tests in PAI-BP subjects in grade VI at SDN 5 Menteng Palangka Raya. The results of the process aspect observation at table 9.

Table 9. Observation Result for the Process Aspect

Observation	Average
Observation 1	3,60
Observation 2	3,60

Observation 3	3,60
Overall Average	**3,60**

In the process aspect, the observation results showed an average score of 3.60, which is in the very good category. This indicates that the implementation of learning outcome tests in PAI-BP learning in grade VI at SDN 5 Menteng Palangka Raya has been carried out systematically and in accordance with the planned procedures. Most of the indicators in the process aspect showed optimal and relatively consistent implementation in each observation, although there were still some indicators with lower scores than others. The results of the questionnaire given to respondents on the process aspect showed a score of table 10.

Table 10. Process Aspect Questionnaire Result

Descriptive Statistics

	N	Mean	Std. Deviation
Mean_Process	32	3.34	.29
Valid N (listwise)	32		

Based on the results of the questionnaire given to respondents on the process aspect, an average score of 3.3438 was obtained with a standard deviation of 0.29. To determine the process evaluation result category, interval criteria were used based on the previous evaluation assessment criteria, namely by calculating the upper and lower limits of the interval using the following formula.

1. $\bar{X} + 1{,}5\ SBx$
   $3{,}34+(1{,}5\times 0{,}29)=3{,}34+0{,}44=3{,}78$
2. $\bar{X} - 1{,}5\ SBx$
   $3{,}34-0{,}44=2{,}90$

Then, the results of the upper and lower limits of the interval were classified into table 11.

Table 11. Process Aspect Evaluation Assessment Criteria

Interval	Criteria
$X \geq 3{,}78$	Very High
$3{,}78 > X \geq 3.34$	High
$3.34 > X \geq 2{,}90$	Low
$X < 2{,}90$	Very Low

Based on the process aspect assessment criteria table above, the average questionnaire score obtained from the process aspect of 3.34 is in the range of 3.78 > X ≥ 3.34, which is classified as high. This finding shows that the consistency of implementation, stakeholder involvement, program progress records and reports, and the accuracy of the target group in the implementation of learning outcome tests conducted in the PAI-BP subject in grade VI at SDN 5 Menteng Palangka Raya have been positively assessed by students.

In the indicator of consistency of implementation, students assessed that the tests were carried out in an orderly manner and in accordance with applicable procedures. The indicator of stakeholder involvement, particularly the role of teachers in managing the implementation of tests, also received a positive assessment. Teachers, as stakeholders, played an active role in ensuring that the learning outcome tests ran smoothly. This is in line with the research by Yusuf & Hamami (2022) hat teachers are required to be able to assess students' abilities in ongoing learning activities. The assessment aims to determine the extent to which learning objectives have been achieved, as well as to provide a basis for continuous assistance in order to improve the quality of education for students at the next level. However, the questionnaire results also show that students have not fully benefited from the program's progress records and reports, indicating that the documentation and

delivery of test results to students can still be improved. An effective learning outcome test process emphasizes not only the implementation of the test but also the systematic recording and reporting of test results as a basis for learning decision-making (Adela et al., 2025).

The indicator of target group accuracy shows that learning outcome tests have been given to students in accordance with the learning targets. However, there are still differences in respondent scores, with students feeling that the tests administered do not fully take into account the individual characteristics of each student. Therefore, improving quality in the process aspect requires guidance that focuses not only on procedural consistency but also on flexibility in the implementation of learning outcome tests oriented towards student needs.

d. Evaluation of the Product Aspects

Product evaluation is carried out to assess the extent to which the desired results have been achieved and are in line with the objectives formulated previously (Rosyiah et al., 2022). Product evaluation also aims to examine the learning outcomes achieved by students after participating in a series of learning activities, as well as to assess the extent to which students are able to participate effectively in the activities that have been carried out (Sardimi et al., 2022). Evaluation of the product aspect in this study analyzes the results or outputs obtained from the learning outcome test, including the achievement of objectives, the benefits obtained by users, and the consistency of sustainability of benefits. The product aspect is analyzed to determine the results or outputs obtained from the learning outcome test, namely the achievement of objectives the benefits obtained by users, and the consistency of sustainability of benefits. Results of product aspect observation on table 12.

Table 12. Observation Results for Product Aspects

Observation	**Average**
Observation 1	2,67
Observation 2	3,00
Observation 3	3,00
Overall Average	**2,89**

The results of observations on the product aspect obtained an average score of 2.89, ranging from 2.51 to 3.25 and falling into the good category. These findings indicate that the implementation of learning evaluation has had a fairly positive impact on the expected outcomes. However, the average score for the product aspect is still lower than that for the input and process aspects, indicating room for improvement in the product aspect in the implementation of learning outcome tests in PAI-BP learning in grade VI at SDN 5 Menteng Palangka Raya. The questionnaire results given to respondents on the product aspect obtained a score of table 13.

Table 13. Product Aspect Questionnaire Result

Descriptive Statistics

	N	Mean	Std. Deviation
Mean_Product	32	3.39	.31
Valid N (listwise)	32		

Based on the results of the questionnaire given to respondents on the product aspect, an average score of 3.39 with a standard deviation of 0.31 was obtained. To determine the category of process evaluation results, interval criteria were used based on the previous evaluation assessment criteria, namely by calculating the upper and lower limits of the interval using the following formula.

1. $\bar{X} + 1{,}5\ SBx$

3,39+(*1,5*×0,31)=3,39+0,47=3,86

2. $\bar{X} - 1{,}5\ SBx$

   3,39−0,47=2,92

Then, the results of the upper and lower limits of the interval were classified into table 14.

Table 14. Product Aspect Evaluation Assessment Criteria

Interval	Criteria
$X \geq 3{,}86$	Very High
$3{,}86 > X \geq 3.39$	High
$3.39 > X \geq 2{,}92$	Low
$X < 2{,}92$	Very Low

Based on the evaluation criteria table for the product aspect above, the average questionnaire score obtained was 3.39, which falls within the range of 3.86 > X ≥ 3.39, thus falling under the high criteria. This finding shows that the achievement of objectives, the benefits obtained by users, and the consistency of the sustainability of benefits in the implementation of learning outcome tests carried out in the PAI-BP subject in grade VI at SDN 5 Menteng Palangka Raya have been able to describe the results of the tests that have been carried out in accordance with the expected objectives.

In terms of the achievement of objectives indicator, students assessed that the test results reflected their understanding of the PAI-BP material that had been studied. This was also in line with the results of observations in the field and documentation of teacher score records during the implementation of the learning outcome test. The indicator of benefits obtained by users showed that the test results were considered useful, especially as a means of determining the level of mastery of the material. However, the results of the questionnaire from the respondents show that the use of test results as learning feedback is not yet fully optimal. Some students still view test results as final scores, not as a basis for learning improvement. This is in line with Huljannah (2021) opinion that test results should be used as a basis for decision making and improvement of the teaching and learning process. Meanwhile, the questionnaire results on the indicator of consistency of sustainability of benefits show that the benefits of test results are not yet fully felt in a sustainable manner. This condition indicates that the follow-up on test results still needs to be improved so that the implementation of learning outcome tests does not only stop at assessing results but continues to improving learning strategies and the quality of student learning (Tazkirah et al., 2024).

The implementation of learning outcome tests in PAI-BP learning, which was analyzed using the CIPP model based on research results, was generally in the good and high categories. These findings indicate that the evaluation of learning outcome tests is not only carried out as a final assessment activity, but is an integral part of the overall learning process. The evaluation carried out in learning is oriented towards decision-making and program implementation, not merely as a measurement of learning outcomes (Qodir, 2017). This is in line with Idrus (2019) opinion that evaluation is positioned to achieve learning objectives as well as a means of reflection for teachers in improving the quality of the learning process.

The findings in this study still have several indicators in aspects that require serious attention. This condition is in line with the message of QS. Al-Hasyr verse 18, which emphasizes the importance of evaluating what has been done as a basis for future improvements. The test results are used as a means of reflection and continuous improvement so that the PAI-BP learning objectives can be achieved optimally.

The findings of this study are in line with the findings of a number of previous studies on PAI learning evaluation using the CIPP model in several aspects that are not yet optimal. The first relevant study was conducted by Anisah Rahmiwati (2022) with the title "Evaluation of PAI Distance Learning with the CIPP Model at SMPIT Asshiddiqqiyyah & SMP Al Wildan Tangerang". This study used a mixed method to measure the effectiveness of online PAI learning. The results showed that the implementation of Distance Learning (PJJ) in both schools was generally running

well. However, when reviewed more deeply based on the aspects of the CIPP model, there was still room for improvement to optimize its implementation. This development needs to focus primarily on the input and process aspects, through innovation in distance learning methods and updates to several supporting components. It is hoped that student learning outcomes at these schools will not only be fulfilled administratively, but will also be able to increase student potential through the design and implementation of more effective and sustainable education programs in each institution (Rahmiwati, 2022).

Dani Darmawan (2021) in his research entitled "Evaluation of the CIPP Model in PAI at MI Muhammadiyah Suruh & MIN 1 Trenggalek" used qualitative methods through observation, interviews, and documentation with the CIPP model approach. The purpose of the study was to determine the effectiveness of PAI learning at MI and MIN based on the CIPP model. The results showed that the evaluation of PAI learning using the CIPP model at MIM Suruh and MIN 1 Trenggalek revealed differences in achievement. At MIM Suruh, the context aspect was considered good in terms of vision, mission, and learning objectives, but improvements were still needed in the learning principles. In terms of input, students' intellectual abilities were classified as good, but creativity, motivation, and discipline still needed to be improved, while teachers needed to strengthen their pedagogical competencies. The process aspect showed that the learning strategies were quite good, but the use of models, learning media, and the completeness of facilities and infrastructure were not yet optimal. In terms of product, the achievements were limited to short-term results, while medium and long-term results had not yet been achieved. Meanwhile, MIN 1 Trenggalek generally shows good results in all aspects of CIPP. However, several indicators still need improvement, especially in the context aspect related to learning principles, the input aspect in student creativity, the process aspect in optimizing learning models and media, and the product aspect in achieving long-term results, particularly madrasah independence (Darmawan, 2020).

When compared to previous studies, the findings of this study are consistent with various previous studies which state that the application of the CIPP model in learning outcome tests is able to provide a comprehensive picture of the quality of the implementation of learning outcome tests. Several studies show that the CIPP model is effective in identifying gaps between planning and implementation of learning evaluation. However, there are also studies that find that the product aspect is often not fully utilized for learning improvement. This difference shows that the contribution of this study lies in emphasizing the importance of continuity between the four aspects of CIPP in learning evaluation.

The findings of this study imply the need to improve teachers' competence in designing and utilizing learning outcome tests more reflectively. Teachers are expected to not only focus on the preparation and implementation of tests, but also on the analysis of test results as a basis for improving learning. This study also reinforces the relevance of the CIPP model as a comprehensive and practical evaluation framework in PAI-BP learning. This model allows researchers to assess learning comprehensively, from context to product in the form of test scores achieved.

Nevertheless, this study is still limited by the small sample size and scope, which was only conducted in one educational unit, so the results cannot be generalized broadly. The research instruments used are still dominated by a quantitative approach, so they do not fully explore the subjective experiences of teachers and students in conducting learning outcome tests in PAI-BP subjects.

Based on the limitations of this study, the researcher suggests that future studies involve a larger sample and a wider scope using a mixed method approach to obtain more in-depth and comprehensive data related to learning outcome tests in PAI-BP subjects. Future research can also focus on utilizing the learning outcome test results from PAI-BP subjects as a basis for educational decision-making and developing more innovative learning strategies. Thus, the findings of this study are expected to contribute more significantly to improving the quality of PAI-BP learning.

CONCLUSION

The implementation of the PAI-BP learning outcome test in grade VI of SDN 5 Menteng Palangka Raya, evaluated using the CIPP model, is generally categorized as high and aligned with

learning objectives. This indicates that the CIPP model is effective in providing a comprehensive evaluation of learning outcome tests, highlighting both strengths and areas for improvement. Overall, the evaluation confirms that the test implementation supports the intended learning goals.

The findings show that all four aspects of CIPP have been implemented according to evaluation standards, although with varying levels of achievement. The context aspect reflects good alignment with learning objectives and student needs. The input aspect indicates adequate readiness of resources and planning, though the use of test results for feedback and follow-up remains suboptimal. The process aspect demonstrates systematic and consistent test implementation, supported by active teacher involvement. Meanwhile, the product aspect shows that the test is able to capture students' understanding, but its utilization for continuous improvement still needs strengthening.

These results emphasize that the integration between planning, implementation, and utilization of test outcomes is crucial for improving the quality of PAI-BP evaluation practices. The CIPP model proves to be a comprehensive framework that supports holistic assessment from context to product. However, the study is limited by its small sample size and focus on a single school, which restricts the generalizability of the findings. In addition, the use of predominantly quantitative instruments limits deeper exploration of teachers' and students' experiences.

Based on these limitations, future research is recommended to involve larger and more diverse samples across different educational settings. The use of mixed methods is also suggested to obtain more comprehensive and in-depth data. Further studies should focus on optimizing the use of test results as a basis for instructional improvement and decision-making, as well as developing more innovative learning strategies in PAI-BP to enhance overall learning quality.

DAFTAR PUSTAKA

Adela, Z. A., Izzah, Y. N., Nuriyah, Z., Akil, A., & Azis, A. (2025). Menakar Keefektifan Tes : Prinsip-Prinsip Kunci dalam Pengembangan Instrumen Evaluasi Pembelajaran. *Hidayah: Cendekia Pendidikan Islam Dan Hukum Syariah*, *2*(2), 114–125. https://doi.org/10.61132/hidayah.v2i2.932

Adinda, A. H., Siahaan, H. E., Raihani, I. F., Aprida, N., Fitri, N., & Suryanda, A. (2021). Penilaian Sumatif dan Penilaian Formatif Pembelajaran Online. *Report Of Biology Education*, *2*(1), 1–10. https://doi.org/10.37150/rebion.v2i1.1024

Akbar, A., Rezki, A., Putri, E., Rahmatika, N. U., Ainin, A., & Nugraha, M. T. (2024). Penerapan Evaluasi Portofolio dalam Pembelajaran Pendidikan Agama Islam dan Budi Pekerti di SMP Negeri 6 Sukadana. *Journal of Education Research*, *5*(4), 5567–5575. https://doi.org/10.37985/jer.v5i4.1832

Andriani, I., Hartati, Z., & Surawan, S. (2021). The CIPP Analysis of PAI Development Program for Mualaf in MCI Palangka Raya. *At-Tarbiyat: Jurnal Pendidikan Islam*, *04*(03), 499–511. https://doi.org/10.37758/jat.v4i3.316

Armedi, R. (2025). Karakteristik Tes yang Baik dan Proses Penyusunan Instrumen Tes untuk Pembelajaran di Sekolah. *PENDAGOGIA: Jurnal Pendidikan Dasar*, *5*(April), 10–17.

Darmawan, D. (2020). *Evaluasi Model Context Input Product (CIPP) Kurikulum 2013 Pada Mata Pelajaran Pendidikan Agama Islam Di MI Muhammadiyah Suruh Dan MIN 1 Trenggalek*. IAIN Kediri.

Fitrianti, L. (2018). Prinsip Kontinuitas dalam Evaluasi Pembelajaran. *Al-Ishlah: Jurnal Pendidikan*, *10*(1), 89–102. http://www.journal.staihubbulwathan.id

Hasmawati, H., & Muktamar, A. (2023). Asesmen dalam Kurikulum Merdeka Perspektif Pendidikan Agama Islam. *Indonesian Journal of Innovation Multidisipliner Research*, *1*(3), 197–211. https://doi.org/10.31004/ijim.v1i3.20

Huljannah, M. (2021). Pentingnya Proses Evaluasi Dalam Pembelajaran Di Sekolah Dasar. *EDUCATOR : Directory of Elementary Education Journal*, *2*(2), 164–180. https://doi.org/10.58176/edu.v2i2.157

Idrus. (2019). Evaluasi dalam Proses Pembelajaran. *Adaara: Jurnal Manajemen Pendidikan Islam*, *9*(2), 920–935. https://doi.org/https://doi.org/10.35673/ajmpi.v9i2.427

Inanna, Rahmatullah, & Hasan, M. (2021). *Evaluasi Pembelajaran: Teori dan Praktek*. Penerbit Tahta Media Group.

Indriani, N., Salsabila, Z. P., Ummah, U. K., & Putri, N. R. (2021). Pengaruh Evaluasi Formatif Terhadap Hasil Belajar Fiqih Peserta Didik Kelas IV MINU Plus Islamiyah Sidoarjo. *Al-Mudarris : Jurnal Ilmiah Pendidikan Islam*, *4*(2), 181–194.

Isnaeni, N., Apriliani, D., & Habibi, B. (2024). Evaluasi Program Literasi Sekolah Menggunakan Model Context Input Process dan Product (CIPP) pada SMA. *Journal of Education Research*, *5*(3), 3245–3252. https://jer.or.id/index.php/jer/article/view/1443

Khaerudin. (2022). *Evaluasi Program Pembelajaran Pesantren*. Trussmedia Grafika.

Khasanah, S., Yahya, M. D., Rindihastuti, A., & Fauziyah, A. N. (2024). Analisis Pengembangan Kurikulum pada Pendidikan Agama Islam. *Journal of Instructional and Development Researches*, *4*(6), 574–586. https://doi.org/10.53621/jider.v4i6.424

Kurniawati, E. W. (2021). Evaluasi Program Pendidikan Perspektif Model CIPP (Context , Input , Process , Product). *Ghaitsa : Islamic Education Journal*, *2*(1), 19–25. https://doi.org/10.62159/ghaitsa.v5i3.168

Laila, I., Pariati, E., & Widyati, E. (2024). Pengembangan Tes-tes Hasil Belajar. *Dharma Acariya Nusantara : Jurnal Pendidikan, Bahasa Dan Budaya*, *2*(2), 12–25. https://doi.org/10.47861/jdan.v2i1.1139

Mardapi, D. (2019). *Pengukuran, Penilaian dan Evaluasi Pendidikan*. Nuha Medika.

Munandar, A., Alfathan, P. I., Pitri, N., Karmisa, N., Azlaini, M., Alfikri, Puspita, C., Nurazizah, Karomah, L., & Amaliyah, E. (2023). Evaluasi Pembelajaran Dalam Meningkatakan Kualitas Pendidikan. *Innovative: Journal Of Social Science Research*, *3*(6), 786–798. https://j-innovative.org/index.php/Innovative/article/view/6329

Munaroh, N. L. (2024). Asesmen dalam Pendidikan : Memahami Konsep,Fungsi dan Penerapannya. *Dewantara: Jurnal Pendidikan Sosial Humaniora*, *3*(3), 281–297. https://doi.org/10.30640/dewantara.v3i3.2915

Oktiviana, N., Nastiti, S. A., & Mulyono, M. (2024). Problematika Evaluasi Pembelajaran Pendidikan Agama Islam di Era Merdeka Belajar. *Jurnal PAI: Jurnal Kajian Pendidikan Agama Islam*, *3*(1), 82–93. https://doi.org/10.33507/pai.v3i1.1990

Pamulatsih, H., & Zulfitria. (2024). Evaluasi Pembelajaran dalam Peningkatan Kompetensi Peserta Didik. *Jurnal Pendidikan: SEROJA*, *3*(2), 65–78.

Pimentel, J. L. (2019). Some Biases in Likert Scaling Usage and its Correction. *International Journal of Sciences: Basic and Applied Research (IJSBAR)*, *45*(1), 183–191. http://gssrr.org/index.php?journal=JournalOfBasicAndApplied

Qodir, A. (2017). *Evaluasi dan Penilaian Pembelajaran* (I). Penerbit K-Media. https://doi.org/10.5346/trbane.1954.193

Rahman, A. A., & Nasryah, C. E. (2019). *Evaluasi Pembelajaran*. Penerbit Uwais Inspirasi Indonesia.

Rahmiwati, A. (2022). *Evaluasi Pembelajaran Jarak Jauh Dengan Menggunakan Model CIPP*. UIN Syarif Hidayatullah.

Ratnawulan, E., & Rusdiana, H. A. (2015). *Evaluasi Pembelajaran*. Pustaka Setia Bandung.

Ridha, A. R., Rahmatullah, N. A., Firdaus, A. N. N., & Wicaksono, Y. (2025). Hasil Belajar Sebagai Objek Penilaian (Kognitif, Afektif, dan Psikomotorik). *Inovasi: Jurnal Ilmiah Pengembangan Pendidikan*, *3*(3), 1–8.

Ropii, M., & Fahrurrozi, M. (2017). Evaluasi Hasil Belajar. In *Yogyakarta: Pustaka Pelajar*. Penerbit Universitas Hamzanwadi Press.

Rosyiah, S. A., Saudah, S., & Jennah, R. (2022). The Evaluation Of CIPP Model In Learning Centers In Islamic Kindergarten Darussalam Palangka Raya. *Al-Hikmah: Indonesian Journal Of Early Childhood Islamic Education*, *6*(1), 83–92. https://doi.org/10.35896/ijecie.v6i1.248

Sapitri, A., & Maryati, M. (2022). Peran Pendidikan Agama Islam Dalam Revitalisasi Pendidikan Karakter. *Al-Afkar,Journal for Islamic Studies*, *5*(1), 252–266. https://doi.org/10.31943/afkarjournal.v5i1.229

Sardimi, Ariani, E., & Aghnaita. (2022). Evaluation of Multi-Literacy Education Program at Community Learning Activity Center in Palangka Raya City. *Tarbawi: Jurnal Keilmuan Manajemen Pendidikan*, *8*(01), 1–14. https://doi.org/10.32678/tarbawi.v8i01.4994

Taqiyuddin, T., Supardi, S., & Lubna, L. (2024). Evaluasi Formatif dan Sumatif dalam Pembelajaran Pendidikan Agama Islam. *Jurnal Ilmiah Profesi Pendidikan*, *9*(3), 1936–1942. https://doi.org/10.29303/jipp.v9i3.2392

Tazkirah, S., Rohani, Purnama, R., & Subyanto, E. (2024). Strategi Remedial dan Pengayaan Sebagai Tindak Lanjut Assesmen Pembelajaran Untuk Meningkatkan Hasil Belajar Siswa SMP Darul Furqan Kota Bima. *Action Research Journal Indonesia (ARJI)*, *6*(4), 472–485. https://doi.org/10.61227/arji.v6i4.239

Tsani, I., Arsyadana, A., Sufirmansyah, & Shafira, E. (2021). Evaluasi Model CIPP Pada Pembelajaran PAI Dan Budi Pekerti DI SMA Negeri 7 Kota Kediri. *Al-Tadzkiyyah: Jurnal Pendidikan Islam*, *12*(1), 17–45. https://doi.org/10.24042/atjpi.v12i1.8165

Tsawab, M. C., Khasanah, U., Afrida, N. F., & Maisyanah. (2019). Evaluasi Pembelajaran Pendidikan Agama Islam dan Implikasinya Terhadap Motivasi Belajar Peserta Didik MTs NU Banat Kudus. *Al-Mudarris : Jurnal Ilmiah Pendidikan Islam*, *2*(2), 93–112. https://doi.org/10.23971/mdr.v3i1.2127

Wan Azman, D. S. (2025). Evaluasi Program Pembelajaran Pendidikan Agama Islam. *JURNAL EDUKASI*, *13*(1), 162–170.

Wildana, W., Zuhdiah, Z., & Darwis, D. (2024). Penerapan Penilaian Formatif Pada Pembelajaran Pendidikan Agama Islam Di SD Negeri 02 Kampung Baru. *Passikola: Jurnal Pendidikan Dasar & Madrasah Ibtidaiyah*, *1*(2), 110–127. https://doi.org/10.46870/passikola.v1i2.1290

Wirawan. (2016). *Evaluasi: Teori, Model, Metodologi, Standar, Aplikasi, dan Profesi* (Cetakan 3). Penerbit PT. Rajagrafindo Persada.

Yuniar, M., Rakhmat, C., & Saepulrohman, A. (2015). Analisis HOTS (High Order Thinking Skills) pada Soal Objektif Tes dalam Mata Pelajaran Ilmu Pengetahuan Sosial (IPS) Kelas V SD Negeri 7 Ciamis. *Pedadidaktika: Jurnal Ilmiah Mahasiswa Pendidikan Guru Sekolah Dasar*, *2*(2), 187–195.

Yuniarti, Y., Abdiyah, L., Nurjanah, S., Siregar, S. L., & Riani, P. (2021). Penelitian Evaluatif dalam Pendidikan. *Yasin: Jurnal Pendidikan Dan Sosial Budaya*, *1*(1), 73–87. https://doi.org/10.58578/yasin.v1i1.14

Yusuf, M., & Hamami, T. (2022). Peran Guru Pendidikan Agama Islam Menyiapkan Peserta Didik dalam Menghadapi Tes Asesmen Kompetensi Minimum. *Jurnal Basicedu*, *6*(2), 3012–3024. https://doi.org/10.31004/basicedu.v6i2.2571