

CIPP Model: Evaluation of the Implementation of Junior High School Physical Education

Wawan Eko Prasetyo¹, Khoirul Fatihin², Duwi Susanti³

¹Physical Education, University of Malang, Indonesia

²Physical Education, University of Surabaya, Indonesia

³Islamic Religious Education, University Islam Darul Ulum, Indonesia

Email author corresponding: wawaneko8837@gmail.com

Abstract:

Background: This research aims to evaluate the improvement of students' knowledge in Physical Education learning (PJOK) at junior high schools using the CIPP evaluation model (Context, Input, Process, Product). **Aims:** Provide data-based recommendations for improving the planning, implementation, and evaluation of PJOK learning to support continuous enhancement of program quality. **Methods:** This research employed a qualitative descriptive approach, with interviews, observations, and questionnaires serving as data collection methods. The study population consisted of 374 students from four schools, using a saturated sampling technique. Overall, the PJOK learning program was implemented in accordance with the standards set by Permendiknas No. 16 of 2007, Permendiknas No. 27 of 2007, and Permendiknas No. 22 of 2016. **Result:** The findings indicate that the PJOK learning programs in both public and private junior high schools across the two districts obtained a total score of 34,221 (65.2%), falling into the "good" category. Based on the CIPP analysis, the average scores were as follows: Context (11,762 or 62.9%, good category), Input (14,713 or 60.5%, good category), Process (5,219 or 69.6%, good category), and Product (2,527 or 67.6%, good category). Data were analyzed using qualitative and quantitative descriptive techniques, supported by inductive analysis grounded in field findings. **Conclusion:** The PJOK learning programs in public and private junior high schools in the Selorejo and Kesamben Districts are well implemented, with all CIPP evaluation components categorized as good, demonstrating successful program implementation in accordance with national standards.

Keywords: *Evaluation CIPP Model, Physical Education*

Introduction

Background of the study: One of the most crucial efforts a nation can undertake to improve the quality of its human resources and achieve the goal of producing an intelligent generation is to invest in education (Botha et al., 2025; Chaka, 2022). The purpose of education is to develop skills, shape national character, and build civilization so that the nation becomes increasingly enlightened and eliminates all forms of ignorance and backwardness (UU RI No. 3 Tahun, 2005). Education and learning cannot be separated from the teacher's role specifically, how teachers apply their knowledge to the material being taught and how they select learning strategies that students find enjoyable to achieve the expected learning outcomes (L, 2019). Whenever an educational program has predetermined objectives, it becomes the obligation of educators to conduct an evaluation. Program evaluation is highly important and valuable, especially for those responsible for making decisions.

Decision-makers can use the results of program evaluations as a foundation for determining next steps for programs that are currently being implemented or have already been completed (Istiyani & Utsman, 2020). Evaluation cannot be separated from the learning process because it is

an integral component of instructional activities (Ahsan et al., 2021; Kane & Broomell, 2022). Conducting assessment as part of educational practice serves an essential purpose, functioning as a measurement tool or method to determine the level of mastery students have achieved based on the learning materials provided, so that learning objectives become accurate and persuasive after evaluation is carried out (Lee, 2019). A systematic investigation that is designed, implemented, and reported to assist clients in determining and/or improving the value and/or benefits of an educational program is referred to as evaluation (Mahmudi, 2011).

Evaluation in the learning process ultimately serves to provide systematic feedback to teachers, enabling them to refine student remediation programs and optimize the overall quality of instruction (Obinna Joseph, 2021; Tustianti et al., 2019). The incorporation of assessment into teaching activities is not optional but fundamental, as it allows educators to obtain comprehensive information across the cognitive, affective, and psychomotor domains of student learning outcomes (Viet & Hanh, 2021; Wang et al., 2022). In the context of physical education, sports, and health, evaluation must be designed and implemented with precision, holism, and strict adherence to the principles of sound assessment to ensure validity, reliability, and fairness. Achieving effective evaluation requires coherent alignment among instructional planning, learning implementation, and assessment procedures (Özer, 2015; Pradina et al., 2020; Webster et al., 2018). These components are inherently interconnected and cannot be separated from one another. Instructional preparation encompassing lesson plans, learning resources, media and equipment, and the selected assessment methods forms the foundation for a functional and effective learning process (Ávila-Moreno et al., 2018; Sharafati et al., 2021).

Literature review: Physical Education (PE) is not solely concerned with physical outcomes, but is increasingly viewed as a pedagogical domain that requires systematic program evaluation to ensure alignment among curriculum goals, resource readiness, instructional implementation, and student learning outcomes. The CIPP model (Context, Input, Process, Product), developed by Stufflebeam, is widely recommended for this purpose because it provides comprehensive and actionable information to support program improvement and accountability rather than merely judging program success or failure. Several studies highlight that CIPP effectively serves as a guide for planning, implementing, and assessing educational programs, as it requires evaluators to consider contextual policy needs, resource availability, instructional dynamics, and expected outcomes. In the context of secondary Physical Education, emerging empirical evidence demonstrates the relevance of CIPP-based evaluation. A quantitative study conducted in junior high schools in Huaibei, China, found significant differences in the implementation of PE programs across institutions, particularly within the Input (availability of resources and instructor quality) and Process (execution of learning activities) dimensions. These two dimensions were identified as the strongest predictors of variation in program quality, indicating a need for policy attention toward strengthening teacher capacity and the organization of physical activity programs to improve PE effectiveness. Recent literature also highlights the growing integration of digital technology in PE, including video analysis tools, learning management systems (LMS), and wearable devices, which have proven beneficial for objective performance monitoring, personalized feedback, and more accurate motor skill assessment.

However, systematic reviews emphasize that although technology enhances engagement and physical activity levels, most studies still lack a holistic evaluation framework that connects technology implementation with multidimensional program evaluation. This gap suggests an opportunity to integrate CIPP with modern digital metrics to produce more comprehensive, evidence-based recommendations for PE program development. Methodologically, prior research

on PE evaluation remains fragmented, with many studies focusing solely on program outcomes without linking them to contextual and input variables. Likewise, multidimensional and longitudinal research remains limited, making it difficult to determine how changes in input or instructional processes influence medium- and long-term learning results. Future studies employing mixed-methods designs, larger and more representative samples (e.g., multiple schools or regions), and digitally supported measurement indicators are expected to enhance the validity and practical value of evaluation findings for junior high school PE programs.

Gap analysis: Although international research on Physical Education at the school level has continued to develop, most studies still focus on measuring student learning outcomes such as motor skill performance, physical activity levels, or academic achievement without providing a comprehensive picture of the systemic effectiveness of instructional programs. Existing evaluations tend to be partial, assessing only specific aspects such as teacher performance, student engagement, or the availability of facilities, without examining the causal relationships between learning objectives, resource readiness, instructional implementation, and learning outcomes. In fact, the CIPP evaluation model (Context, Input, Process, Product) offers an ideal framework for assessing these four dimensions simultaneously; however, its application in evaluating Junior High School Physical Education remains highly limited and insufficiently documented, particularly within the Indonesian context. Furthermore, only a few studies integrate modern evaluation technologies such as video analysis, learning management systems, or digital performance-based data as part of both process and product assessment. The absence of a holistic, multimodal, and contextual evaluation approach results in program improvement recommendations that are not fully accurate in addressing the real needs of schools. Therefore, a comprehensive evaluation of the implementation of Physical Education using the CIPP model is needed to generate critical findings that can serve as a foundation for decision-making, instructional quality improvement, and sustainable policy development.

The novelty of this research lies in the comprehensive application of the CIPP (Context, Input, Process, Product) evaluation model to assess the implementation of Physical Education at the junior high school level. Unlike previous research that examined only isolated aspects such as students' motor skill outcomes, teacher performance, or the availability of facilities, this study evaluates the structural interconnectedness and causal relationships between curriculum objectives, resource readiness, instructional implementation, and learning outcomes. In addition, this research integrates technology-based evaluation tools, such as digital learning platforms and performance observation metrics, to strengthen the validity and reliability of both process and outcome measurements. The findings are expected not only to map the strengths and weaknesses of the implementation of Physical Education but also to provide evidence-based strategic recommendations that can be directly utilised for policy decision-making, school quality improvement, and sustained instructional innovation.

Rationale of the study: The evaluation of Physical Education (PE) programs in junior high schools is becoming increasingly essential as educational systems worldwide emphasize competency-based learning, holistic student development, and evidence-based decision-making. Despite the strategic role of PE in developing students' physical literacy, health awareness, motor competence, and psychosocial skills, many schools still implement PE without a systematic mechanism to assess whether instructional objectives, resource readiness, and learning processes are aligned in producing meaningful learning outcomes. Current research trends primarily focus on isolated aspects such as motor skill performance, teacher effectiveness, or facility availability without examining the broader structural and contextual factors that influence program quality.

This fragmentation limits the ability of schools and policymakers to identify root causes of underperformance and to design sustainable improvements.

The CIPP (Context, Input, Process, Product) evaluation model offers a comprehensive framework capable of assessing educational programs holistically by analyzing curriculum relevance, resource adequacy, instructional implementation, and learning outcomes simultaneously. However, its application in junior high school Physical Education particularly within the Indonesian context remains limited. Most existing studies rely on summative outcome measurements and fail to integrate diagnostic information that can guide program refinement. Additionally, the rapid emergence of digital technologies such as video analysis tools, learning management systems, and digital performance-based assessments has not yet been optimally utilized within PE program evaluation. This creates a methodological gap between modern instructional practices and the evaluation tools used to assess their effectiveness.

Given these limitations, a rigorous and comprehensive evaluation using the CIPP model is necessary to generate accurate, actionable insights regarding the implementation of Physical Education in junior high schools. Such an evaluation is expected to provide a clearer understanding of how contextual factors, resource readiness, instructional quality, and technological integration collectively shape student learning outcomes. Ultimately, the rationale for this study lies in its potential to support data-driven decision-making, improve the quality and equity of PE delivery, and contribute to the development of more responsive and future-oriented Physical Education programs aligned with current educational demands and technological advancements.

Material & Methods

Research Design: This research employed the 4D development model originally introduced by Borg and Gall (2003) and later refined by Thiagarajan, as cited in (Elvera and Yesita Astarina, 2021). The model consists of four systematic stages: Define, Design, Develop, and Disseminate. In the Define stage, the researchers identified core problems, clarified learning needs, and determined the objectives underlying the product to be developed. The Design stage focused on creating the initial framework, determining content structure, instructional strategies, and preparing prototype drafts. The Develop stage involved expert validation, product refinement, and field testing to ensure feasibility, accuracy, and usability. Finally, the Disseminate stage emphasized broader implementation, including distribution, limited-scale trials, and evaluation of user acceptance. This structured 4D approach ensured that the development process proceeded logically, iteratively, and met the criteria of a scientifically grounded research and development methodology.

Participant: The population of this research consisted of 1,500 students enrolled in four junior high schools in the Selorejo District. A purposive sampling technique was applied to ensure that participants met the specific requirements of the study. Inclusion criteria required students to be officially registered as active learners, fall within the typical junior high school age range, and regularly participate in physical education classes, while exclusion criteria eliminated students with medical limitations, those absent during data collection, or those who did not provide consent. Based on these criteria, a total of 511 students were selected as the final sample. The use of purposive sampling ensured that the participants were highly relevant to the study objectives, thereby enhancing the validity and representativeness of the data collected.

Instrument: Interview guidelines, documentation sheets, and observation forms served as the primary data-collection instruments in this study, which evaluated the Physical Education, Sports, and Health program using the CIPP model among junior high school students in the Selorejo and Kesamben districts. Each instrument was first reviewed and assessed by Prof. M. E. Win through an expert-judgment validation process to ensure its relevance, clarity, and alignment with the predetermined criteria for each CIPP component. This validation step was essential to establish the content validity and methodological rigour of the instruments before field implementation. The blueprint and operational structure of the instruments, organised according to the required components of the evaluation model.

Table 1. Questionnaire Instrument Blueprints for Students

Concept	No.	Variable	Indicator	Total	Item Number
CIPP Evaluation	1	Context	Curriculum	1	1
			Benefits of Physical Education learning	7	2, 3, 4, 5,6, 7, 11
			Motivation to learn Physical Education	2	8, 9,
	2	Input	Teacher background	4	12,13,14,15
			Teaching methods	3	22, 25, 26
			Teaching materials	1	20
			Physical Education learning media	1	10
			Physical Education facilities and infrastructure	4	16, 17, 18, 19
	3	Process	Classroom management by Physical Education teacher (introduction, motivation, examples, instructions) Understanding of Physical Education material	4	21, 23, 24, 27
	4	Product	Ability to practice Physical Education	1	28
			Curriculum	1	29
			Item Total		

The methodology employed in this research integrates both qualitative and quantitative data analysis as part of the evaluation of the learning program. Based on the data collected, an inductive qualitative analysis approach was applied to formulate emerging hypotheses. This process separates qualitative data analysis into two main stages: pre-field analysis, conducted before data collection to refine research focus and prepare analytical guidelines, and in-field analysis, carried out during data collection to interpret findings as they unfold in real conditions. This analytical flow aligns with the qualitative data analysis framework proposed by (Sugiyono, 2010), emphasising continuous interaction between data reduction, data display, and conclusion drawing. To illustrate the analytical process clearly, the study adopts the Flow Model, which outlines the

core components of qualitative data analysis and highlights how each stage informs subsequent interpretation. The model provides a structured yet flexible framework, ensuring that data interpretation remains grounded in field realities while allowing emerging patterns to shape the final conclusions.

Table 2. Components of Data Analysis (Flow Model)

Anticipation	Data Collection Period		
	Data Reduction		
	During	After	Analysis
	Data Display		
	During	After	Analysis
	Drawing/Verification		
	Buring	After	Analysis

Analysis plan: The qualitative data were processed using an interactive and continuous procedure until the analysis was completed, following the key components of data analysis described by Sugiyono (2010). This approach ensures that the data remain clear, structured, and interpretable throughout the research process. The analysis includes three essential stages: data reduction, where irrelevant information is filtered and categorized; data display, which involves organizing the data into visual or narrative forms to enhance comprehension; and conclusion drawing and verification, where emerging findings are interpreted and validated. In this study, qualitative data obtained from interviews and narrative responses from physical education (PJOK) teachers at both public and private junior high schools in the Selorejo and Kasemben districts were analyzed using this qualitative data analysis model. Meanwhile, descriptive percentage statistics were employed to analyze the quantitative data collected in the study. This method allows researchers to present numerical findings clearly and systematically. The formula used for percentage-based descriptive analysis is as follows:

$$P = \frac{f}{n} \times 100\%$$

P : percentage

F : total score obtained from all respondents

N : maximum possible score from all respondents

The collected data were categorised according to their respective classifications. Numerical data were classified as quantitative data. These quantitative data were gathered through a questionnaire developed using a Likert scale. The percentage score (P) was calculated by comparing the total score obtained from all respondents (F) with the maximum possible score (N). This calculation allowed the researchers to determine the overall tendency of responses and measure the level of agreement, perception, or attitude represented in each item. The use of the Likert scale ensured that respondents' opinions could be quantified systematically, providing measurable indicators that could be analysed statistically. This approach not only facilitated the interpretation of data trends but also enhanced the reliability and consistency of the measurement process within the research.

Tabel 3. Interpretasi Skor Persentase

Skor Persentase (%)	Interpretasi
0% – 24,99%	Poor
25% – 49,99%	Not Bad
50% – 74,99%	Good
75% – 100%	Verry Good

Results

Results and Discussion: The primary purpose of program evaluation is to ensure the relevance and achievement of objectives, as well as to assess the efficiency, effectiveness, impact, and sustainability of the program. A well-designed evaluation must provide accurate, meaningful, and actionable information that supports sound decision-making. Program evaluation is defined as a systematic and objective assessment of an ongoing or completed program, policy, or intervention, examining both its implementation design and its outcomes.

Within the evaluation process, the initial step involves examining the existing situation to determine what actions are required. This assessment identifies and analyzes the needs that guide program development. Context evaluation focuses on the circumstances, background conditions, and environmental factors that shape the formulation of learning goals and strategies. The primary aim of context evaluation is to identify the strengths and weaknesses of the program or system being evaluated. By understanding these strengths and limitations, evaluators can provide clear direction for necessary improvements.

This evaluation framework supports decision-making by clarifying program needs, ensuring alignment between objectives and implementation, and offering a structured foundation for formulating or refining program goals. Ultimately, context evaluation helps ensure that the program remains relevant, responsive, and capable of achieving its intended impact.

Table 4. Output Descriptive Statistics Variable Contexts

Indicator	Item	Skor		Mean	(%)	Desc.
		Result	Max			
Contexts	The material refers to the Core Competencies (KI) and Basic Competencies (KD) contained in the K-13 curriculum.	11762	18700	3,14	62,9%	Good
	Does the Physical Education (PJOK) teacher explain the connection between previous learning and the lesson to be delivered?	11762	18700	3,14	62,9%	Good
	Does the teacher explain the basic competencies that must be achieved for each new learning material?	11762	18700	3,14	62,9%	Good

Indicator	Item	Skor		Mean	(%)	Desc.
		Result	Max			
	Does the teacher present the material, objectives, benefits, and learning steps to be carried out in each meeting?	11762	18700	3,14	62,9%	Good
	Does the teacher provide warm-up activities that are related to the learning material?	11762	18700	3,14	62,9%	Good
	When delivering the material, does the teacher give examples that relate to students' daily lives?	11762	18700	3,14	62,9%	Good
	I feel that there is no benefit from learning Physical Education (PE).	11762	18700	3,14	62,9%	Good
	I understand that I must participate seriously in PE classes.	11762	18700	3,14	62,9%	Good
	I want to have a healthy and fit body so that I always feel enthusiastic when taking PE lessons.	11762	18700	3,14	62,9%	Good
	The PE subject provides me with knowledge and understanding about motor learning and health.	11762	18700	3,14	62,9%	Good
Mean District					62,9%	Good

Tabel 5. Output Descriptive Statistics Variabel Input

Indicator	Item	Skor		Mean	(%)	Desc.
		Result	Max			
Input	How proficient is the PE teacher in mastering the material presented during PJOK learning?	14713	24310	3,03	60,5%	Good
	In your opinion, is the PE teacher able to properly demonstrate the PJOK movement skills?	14713	24310	3,03	60,5%	Good
	In my view, the PE teacher possesses a high level of competence in teaching PJOK.	14713	24310	3,03	60,5%	Good
	I am aware that my PE teacher's educational background aligns with the subject they teach (PJOK).	14713	24310	3,03	60,5%	Good
	The teacher provides PJOK learning activities that make students feel happy and engaged.	14713	24310	3,03	60,5%	Good
	The teacher's teaching method makes me feel bored.	14713	24310	3,03	60,5%	Good
	The teacher is disciplined in carrying out PJOK learning activities.	14713	24310	3,03	60,5%	Good
	To increase my knowledge of PJOK learning material, I use (1) the textbook, (2) the module, and (3) the student workbook (LKS) for PJOK.	14713	24310	3,03	60,5%	Good
	My teacher sometimes uses videos, animations, and other engaging learning media when teaching PJOK.	14713	24310	3,03	60,5%	Good
	In your opinion, what percentage of PJOK facilities is available in accordance with the learning material?	14713	24310	3,03	60,5%	Good
	I perform physical activities using adequate and supportive facilities and equipment.	14713	24310	3,03	60,5%	Good
	The existing facilities do not help me in carrying out PJOK learning activities.	14713	24310	3,03	60,5%	Good

	The available facilities and equipment for PJOK learning are insufficient.	14713	24310	3,03	60,5%	Good
Mean District					60,5%	Good

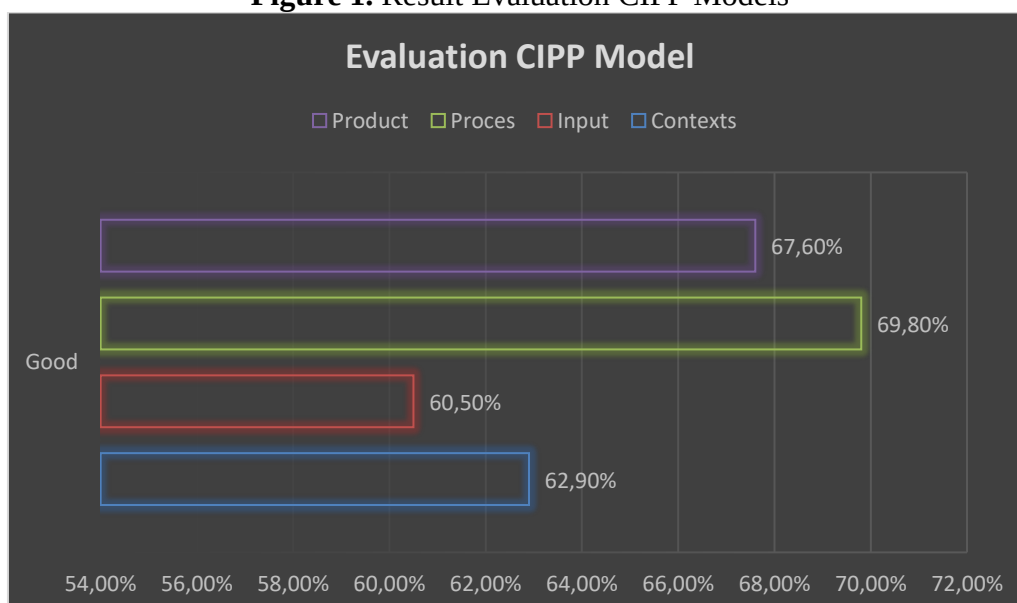
Tabel 6. Output Descriptive Statistics Variabel Proses

Indicator	Item	Total	Skor Max	Mean	(%)	Desc
Proses	At the beginning of the lesson, the teacher greets the students as the class begins.	1421	7480	3,49	69,8%	Good
	The teacher motivates the Physical Education (PE) learning process.	1207	7480	3,49	69,8%	Good
	During each lesson, the PE teacher demonstrates the learning material.	1287	7480	3,49	69,8%	Good
	The teacher gives guidance throughout the learning activities.	1304	7480	3,49	69,8%	Good
Mean District					69,8%	Good

Tabel 7. Output Descriptive Statistics Variabel Product

Indicator	Variable	Total	Skor Hasil	Score Max	Mean	Persen (%)	Ket.
Product	Students are able to understand the material delivered by the Physical Education teacher during the lesson.	1279	2527	3740	3,38	67,6%	Good
	Students are able to perform the physical education motor skills practiced during the learning process.	1248	2527	3740	3,38	67,6%	Good
Mean District						67,6%	Good

Figure 1. Result Evaluation CIPP Models



Discussion

The evaluation of the *Context* component in the curriculum indicator was supported by interview data. From the curriculum perspective, interviews revealed that all four schools involved in the study implemented the 2013 Curriculum (K-13). When PJOK teachers were asked, “*Can the KI and KD materials in the K-13 curriculum be implemented?*” 75% (three teachers) stated that the materials could be delivered at a level of 76–100%. This high percentage indicates strong professionalism among PJOK teachers in integrating curriculum content into learning materials. Curriculum content plays a crucial role and must be properly prepared to ensure learning objectives are achieved. These objectives must align with the Basic Competencies and Competency Standards that students are required to meet, meaning that instructional content must directly support the achievement of competencies and indicators (Arifin, 2015).

For the infrastructure indicator, the statement “*Existing facilities do not support me in conducting PJOK learning activities*” received a score of 662, ranking fifth. Another item, “*Existing facilities support me in carrying out PJOK learning activities,*” ranked fifth as well with a score of 672. These findings were reinforced by interviews with students and PJOK teachers. When asked, “*What percentage of PJOK facilities match the learning material?*”, 60.5% responded that facility availability matched 76–100% of learning needs, while 39.5% estimated availability at 51–75%.

The indicators regarding teacher background and teaching methods showed the highest scores in the input variable based on descriptive statistical analysis. Documentation data confirmed that, among the four teachers sampled, only one had an academic background in PJOK, while the remaining three came from non-PJOK backgrounds, although all held bachelor’s degrees. In terms of teaching experience, two teachers had taught for more than five years, while the other two had taught for one to three years.

Analysis results showed that the indicator *Providing instructions and practice* was crucial in PJOK learning. This is reflected in two items within this indicator receiving the highest overall scores. The item “*The teacher provides instruction during learning*” received the highest total score of 1421. Similarly, the item “*During every lesson, the PJOK teacher demonstrates the learning material*” obtained an average score of 3.49, indicating that students perceive instructional guidance and demonstration as essential components of PJOK learning.

Process evaluation within the CIPP model aims to determine the extent to which program activities are implemented according to plan. It also functions to detect or predict implementation procedures during execution, provide information for program decisions, and serve as a record of operational procedures. The process evaluation in this study focused on the implementation of classroom management by PJOK teachers (introduction, motivation, demonstration, and instruction) in public and private junior high schools in Selorejo and Kesamben.

Results further confirmed that *providing instructions and practice* was the most significant process indicator. The items “*The teacher provides instruction during the learning process*” and “*The PJOK teacher always provides demonstrations during lessons*” both achieved an average score of 3.49 (69.8%), demonstrating that these aspects are viewed by students as essential components of effective PJOK instruction.

Product evaluation aims to measure, interpret, and assess program achievement. This evaluation compares established objectives with actual outcomes through quantitative or qualitative data, such as test scores, percentages, observational data, diagrams, or sociometric information.

Questionnaire analysis showed that the product evaluation identified *Students' ability to understand the material* as the most important indicator in PJOK learning across the four schools, with the corresponding item scoring 1279. The next indicator, *Ability to perform PJOK activities*, scored 1248, showing only a small difference of 31 points. This suggests that both cognitive understanding and practical performance are equally important in the product variable. To strengthen the questionnaire data, researchers also conducted interviews with four PJOK teachers. The results showed that 50% rated the overall implementation of PJOK learning in junior high schools in Kesamben and Selorejo as “adequate,” while the remaining 50% rated it as “highly satisfactory.”

Conclusions

Based on the results of the study, it can be concluded that the Physical Education, Sports, and Health (PJOK) program assessed using the Context, Input, Process, Product (CIPP) model at junior high schools in Selorejo District and Kesamben District achieved an overall score of 34,221, equivalent to 65.2%, which falls into the good category. All four components context, input, process, and product were also rated in the good category. Specifically, the context component obtained a score of 11,762, the input component received 14,713, the process component scored 5,219, and the product component achieved 2,527. These results indicate that the PJOK program is generally well-implemented, with each aspect performing effectively and meeting the expected educational standards.

Acknowledgment

The authors would like to express their sincere gratitude to the principals, teachers, and students of the junior high schools in Selorejo District and Kesamben District for their invaluable participation in this study. Special appreciation is extended to the Physical Education (PJOK) teachers who provided essential data through interviews and questionnaires, contributing greatly to the evaluation of the program using the CIPP model. The authors also acknowledge the support of the school administrations for granting permission to conduct the research and facilitating access to the required documents and learning environments. Lastly, the authors wish to thank all individuals and institutions whose assistance, feedback, and cooperation made this study possible.

Author Contribution Statement

All authors contributed substantially to this research. RS was responsible for formulating the research idea, designing the methodology, and developing the research instruments. MAP was involved in data collection and processing, as well as conducting statistical analyses of the research findings. MA contributed to the theoretical review, interpretation of results, and preparation of the discussion and conclusions. All authors actively participated in the writing, review, and final approval of the manuscript before publication.

AI Disclosure Statement

Part of the manuscript writing process involved the use of artificial intelligence (AI) technology, specifically ChatGPT (OpenAI, GPT-5), which was utilized in a limited manner to assist with the

construction of academic sentence structures, language editing, and English grammar refinement to align with scientific standards. All ideas, data analyses, result interpretations, and research conclusions are entirely the work of the researcher, without any intervention or decision-making from the AI system.

Conflicts Of Interest

The author declares no conflict of interest that could influence the results, interpretation, or writing of this article. The entire research process from planning, data collection, and analysis to report preparation was conducted independently without any external intervention from institutions, sponsors, or individuals. Furthermore, the author did not receive any financial, material, or other forms of support that could potentially introduce bias into the study's findings. Therefore, all results presented in this article fully reflect an objective scientific investigation based on empirical data.

References

- Agustin, A. N., & Kurniawan, A. W. (2021). Pengembangan Media Pembelajaran Variasi Permainan Senam Lantai Berbasis Aplikasi Articulate Storyline. *Sport Science and Health*, 3(6), 369–380. <https://doi.org/10.17977/um062v3i62021p369-380>
- Ahsan, M., Mashuri, M., & Khusna, H. (2021). *Evaluation of laney p ' chart performance Evaluation of Laney p ' Chart Performance*. (April). <http://www.ripublication.com>
- Aminah, A. (2020). Upaya Meningkatkan Prestasi Belajar Pendidikan Agama Islam Dengan Menerapkan Model Pengajaran Kolaborasi Pada Siswa Kelas III Tahun Pelajaran 2016/2017. *JUPE : Jurnal Pendidikan Mandala*, 5(5). <https://doi.org/10.36312/jupe.v5i5.1344>
- Arifin, M. S. (2015). Pengembangan Materi Pembelajaran Makassar. *Education Sience and Technology*, 1(1), 3. <http://ojs.unm.ac.id/index.php/JEST>
- Ávila-Moreno, F. M., Chiroso-Ríos, L. J., Ureña-Espá, A., Lozano-Jarque, D., & Ulloa-Díaz, D. (2018). Evaluation of tactical performance in invasion team sports: a systematic review. *International Journal of Performance Analysis in Sport*, 18(2), 195–216. <https://doi.org/10.1080/24748668.2018.1460054>
- Botha, L., Viviers, H. A., Fouché, J. P., & Kruger, P. (2025). Success factors for postgraduate accountancy and related business students in higher education: a systematic review. *Journal of Further and Higher Education*, 49(6), 749–765. <https://doi.org/10.1080/0309877X.2025.2495957>
- Chaka, C. (2022). Is Education 4.0 a Sufficient Innovative, and Disruptive Educational Trend to Promote Sustainable Open Education for Higher Education Institutions? A Review of Literature Trends. *Frontiers in Education*, 7. <https://doi.org/10.3389/feduc.2022.824976>
- Chukwurah, L. N., Abbah, O. I., Iweama, C. N., Ogugua, J. E., & Ameh, J. (2020). Students' achievement in physical and health education: Effect of discussion teaching method. *International Journal of Human Movement and Sports Sciences*, 8(3), 86–90. <https://doi.org/10.13189/saj.2020.080302>
- Ciris, V. (2020). Investigation of Prospective Teachers' Attitudes towards Game and Physical Activities Course. *Higher Education Studies*, 10(4), 94. <https://doi.org/10.5539/hes.v10n4p94>
- Elvera, & Yesita Astarina. (2021). *Metodologi Penelitian*. Penerbit Andi. <https://books.google.co.id/books?id=p2rDEAAQBAJ>
- Eraut, M. (1982). Evaluation a systematic approach. In *International Journal of Educational Development* (Vol. 2, Issue 3, p. 291). [https://doi.org/10.1016/0738-0593\(82\)90010-4](https://doi.org/10.1016/0738-0593(82)90010-4)
- Harding, L. (2012). Pronunciation Assessment. *The Encyclopedia of Applied Linguistics*, 15(4),

- 57–84. <https://doi.org/10.1002/9781405198431.wbeal0966>
- Irianto, T. (2018). Implementasi Kurikulum 2013 dalam Pembelajaran Pendidikan Jasmani Olahraga dan Kesehatan di Sekolah Dasar. *Jurnal Multilateral*, 13(1), 58–61. <https://ppjp.ulm.ac.id/journal/index.php/multilateralpjk/article/download/6107/5017>
- Istiyani, N. M., & Utsman, U. (2020). Evaluasi Program Model CIPP Pada Pelatihan Menjahit Di LKP Kartika Bawen. *Learning Community: Jurnal Pendidikan Luar Sekolah*, 3(2), 6. <https://doi.org/10.19184/jlc.v3i2.16810>
- Kane, P. B., & Broomell, S. B. (2022). Investigating lay evaluations of models. *Thinking and Reasoning*, 28(4), 569–604. <https://doi.org/10.1080/13546783.2021.1999327>
- Kim, O.-J. (2019). A Study On The Measures For Managing The Quality of Curriculum of Early Childhood Education Department in College With The Application of CIPP Model Based on PDCA. *Journal of the Korea Convergence Society*, 10(1), 215–226. <https://www.koreascience.kr/article/JAKO201909258120684.view>
- L, I. (2019). Evaluasi Dalam Proses Pembelajaran Idrus L 1. *Evaluasi Dalam Proses Pembelajaran*, 9(2), 920–935. <https://doi.org/10.35673/ajmpi.v9i2.427>
- Lee, S. Y., Lee, S. H., & Shin, J. S. (2019). Evaluation of medical humanities course in college of medicine using the context, input, process, and product evaluation model. *Journal of Korean Medical Science*, 34(22), 1–13. <https://doi.org/10.3346/jkms.2019.34.e163>
- Lee, S. young, Shin, J. S., & Lee, S. H. (2019). How to execute Context, Input, Process, and Product evaluation model in medical health education. *Journal of Educational Evaluation for Health Professions*, 16, 1–8. <https://doi.org/10.3352/JEEHP.2019.16.40>
- Mahmudi, I. (2011). CIPP: Suatu model evaluasi program. *Jurnal University of Darussalam*, 6(1), 111–125. <http://doi.org/10.21111/at-tadib.v6i1.551>
- Marsden, J. (1991). Evaluation: Towards a definition and statement of purpose. *Australasian Journal of Educational Technology*, 7(1). <https://doi.org/10.14742/ajet.2276>
- Mufidah, N., Arafat, Y., & Puspita, Y. (2021). The Effect of Training and Teaching Experience on Teacher's Performance. *Proceedings of the International Conference on Education Universitas PGRI Palembang (INCoEPP 2021)*, 565, 5551–5555. <https://doi.org/10.2991/assehr.k.210716.030>
- Munna, A. S., & Kalam, M. A. (2021). Teaching and learning process to enhance teaching effectiveness: literature review. *International Journal of Humanities and Innovation (IJHI)*, 4(1), 1–4. <https://doi.org/10.33750/ijhi.v4i1.102>
- OECD. (2016). PISA 2015 Results (Volume II): Policies and Practices for Successful Schools, PISA. In *OECD Publishing: Vol. II*. <http://dx.doi.org/10.1787/9789264267510-en>
- Obinna Joseph, O. (2021). A Review: Models of Curriculum Evaluation. *Department of Education Foundation Faculty of Education School of Post Graduate Studies Federal University, Dutsin-Ma Katsina State*, 1–34. <https://doi.org/10.14221/AJTE.1988V13N2.1>
- Özer, U. (2015). Evaluation Of Distance Education Process In Sports Higher Education. *ISSN 1822-7864 (Print) ISSN 2538-7111 (Online)* <https://doi.org/10.33225/Pec/22.80.195,7864>
- Pradina, G., Prastuti, I., Rustiadi, T., & Priyono, B. (2020). Evaluation of performance program in tirta wahana swimming club, Grobogan regency. *Journal of Physical Education and Sports*, 9(2), 175–179. <https://doi.org/10.15294/jpes.v9i2.34201>
- Ramadhani, D., Mahardika, I. M. S., & Indahwati, N. (2020). Evaluasi Pembelajaran Pjok Berbasis Daring Terhadap Tingkat Pemahaman Dan Motivasi Belajar Siswa Kelas Iv - Vi Sd Negeri Betro, Sedati - Sidoarjo. *Jurnal Ilmiah Mandala Education*, 7(1), 328–338. <https://doi.org/10.36312/jime.v7i1.1817>

- Rasam, F., Sari, A. I. C., & Karlina, E. (2019). Peran Kompetensi Dalam Meningkatkan Kinerja Guru Di Sma Jakarta Selatan. *Research and Development Journal of Education*, 6(1), 41. <https://doi.org/10.30998/rdje.v6i1.4371>
- Sari, Z. O. (2021). Evaluasi Pelaksanaan Pembelajaran Pendidikan Jasmani Olahraga Dan Kesehatan (Pjok) Saat Pandemi Covid-19 Di Smp Negeri Se-Kabupaten Purbalingga. In *Analisis Standar Pelayanan Minimal Pada Instalasi Rawat Jalan di RSUD Kota Semarang* (Vol. 3). <https://eprints.uny.ac.id/71497/>
- Sharafati, A., Masoud, H., Tiwari, N. K., Bhagat, S. K., Al-Ansari, N., Chau, K. W., & Yaseen, Z. M. (2021). Performance evaluation of sediment ejector efficiency using hybrid neuro-fuzzy models. *Engineering Applications of Computational Fluid Mechanics*, 15(1), 627–643. <https://doi.org/10.1080/19942060.2021.1893224>
- Sugiyono. (2010). *Metode Penelitian Pendidikan: Pendekatan Kuantitatif, Kualitatif, dan R&D*. Alfabeta.
- Suharto. (2015). Evaluasi pelaksanaan kegiatan uji sertifikasi kompetensi keahlian administrasi perkantoran. *Prosiding Seminar Nasional Pendidikan Administrasi Perkantoran, September*, 37–47. <https://jurnal.uns.ac.id/snpap/article/view/26528>
- Sunhaji, S. (1970). Konsep Manajemen Kelas Dan Implikasinya Dalam Pembelajaran. *Jurnal Kependidikan*, 2(2), 30–46. <https://doi.org/10.24090/jk.v2i2.551>
- Suwarno, S. M., & Ismanto, B. (2020). Evaluasi Tempat Uji Kompetensi Teknisi Otomotif dalam Peningkatan Mutu Lulusan Sekolah Menengah Kejuruan. *Kelola: Jurnal Manajemen Pendidikan*, 7(1), 98–109. <https://doi.org/10.24246/j.jk.2020.v7.i1.p98-109>
- Tustianti, R., Nasuka, N., & Sulaiman, S. (2019). Evaluation of gymnastics sports development in pati regency. *Journal of Physical Education and Sports*, 8(4), 71–75. <https://doi.org/10.17977/um065.v4.i3.2024.9>
- Undang-Undang Republik Indonesia Nomor 3 Tahun 2005. (2005). Undang-Undang Republik Indonesia Nomor 3 Tahun 2005 Tentang Sistem Keolahragaan Nasional. *Materia Japan*, 5(1), 1–12. http://joi.jlc.jst.go.jp/JST.Journalarchive/materia1994/44.24?from=CrossRef%250Ahttps://www.bertelsmann-stiftung.de/fileadmin/files/BSt/Publikationen/GrauePublikationen/MT_Globalization_Report_2018.pdf%250Ahttp://eprints.lse.ac.uk/43447/1/India_globalisation%25
- Viet, N. M., & Hanh, N. D. (2021). Assessment Perspectives on the Sports Values and Role of the Higher Education Institution Head for Sports Development in the School. *Higher Education Studies; Vol. 11, No. 4; 2021 ISSN 1925-4741 E-ISSN 1925-475X Published by Canadian Center of Science and Education*, 11(4), 2–7. <https://doi.org/10.5539/hes.v11n4p1>
- Wang, B., Lin, Q., Jiang, T., Yin, H., Zhou, J., Sun, J., Wang, D., & Dai, R. (2022). Evaluation of linear, nonlinear and ensemble machine learning models for landslide susceptibility assessment in southwest China. *Geocarto International*, 0(0). <https://doi.org/10.1080/10106049.2022.2152493>
- Webster, J., Exley, J., Copestake, J., Davies, R., & Hargreaves, J. (2018). Timely evaluation in international development. *Journal of Development Effectiveness*, 10(4), 482–508. <https://doi.org/10.1080/19439342.2018.1543345>
- Wanzer, D. L. (2021). What Is Evaluation?: Perspectives of How Evaluation Differs (or Not) From Research. *American Journal of Evaluation*, 42(1), 28–46. <https://doi.org/10.1177/1098214020920710>