Vol. 14, No. 1, pp. 132-144, 2024 DOI: 10.23960/jpp.v14.i1.202410

Jurnal Pendidikan Progresif

e-ISSN: 2550-1313 | p-ISSN: 2087-9849

http://jurnal.fkip.unila.ac.id/index.php/jpp/

Bilingual Learning Program in Math and Science Subject for Junior High School: An Evaluation using CIPP Model

Rafinda Putri Mentari, Amat Jaedun, Puspa Widyastuti* & Rina Safitri

Department of Research and Evaluation Education, Universitas Negeri Yogyakarta, Indonesia

*Corresponding email: puspawidyastuti.2021@student.uny.ac.id

Received: 17 January 2024 *Accepted: 16 March 2024* *Published: 26 April 2024*

Abstract: Bilingual Learning Program in Math and Science Subject for Junior High School: An Evaluation using CIPP Model. **Objective:** The current research was to reveal the implementation of bilingual learning programs at private junior high schools in D.I. Yogyakarta in terms of context, input, process, product. **Methods:** A descriptive quantitative approach was implemented to 52 students that was chosen by purposive sampling. Data collection used questionnaires, observations and interviews that was analyzed in terms of content validity, construct validity, reliability, and percentage of success. **Findings:** The context component is in accordance with the school's vision, mission, goals and curriculum, the input component of Budi Mulia Dua Middle School with a percentage of 74% is in the good category, Mutiara Persada Middle School with a percentage of 62% is in the good category, Growing Middle School with a percentage of 78% is in the good category, Component process Budi Mulia Dua Middle School with a percentage of 70% in the good category, Mutiara Persada Middle School with a percentage of 77.5% in the good category, Growing Middle School with a percentage of 79% in the good category. The product component shows that the results of the bilingual class are higher than the national class. **Conclusion:** Evaluation of the CIPP model bilingual learning program has gone well. This can be seen from the results of the context, input, process and product components.

Keywords: bilingual learning program, CIPP evaluation model, junior high school students.

To cite this article:

Mentari, R. P., Jaedun, A., Widyastuti, P., & Safitri, R. (2024). Bilingual Learning Program in Math and Science Subject for Junior High School: An Evaluation using CIPP Model. *Jurnal Pendidikan Progresif*, *14*(1), 132-144. doi: 10.23960/jpp.v14.i1.202410.

INTRODUCTION

The use of language has an urgency where the scope is not only mastery of foreign languages, especially English. Increasing language resources through learning can provide provisions for students to access updated knowledge and advances in information in the world (Uccelli, 2023). In addition, students can improve good interaction and communication skills (Tulomana, Tagimaucia, & Chand, 2023). Human resources in language can be developed, one of which is through bilingual programs in schools and implementing an integrative curriculum. An integrative curriculum is a means to support integrated learning to strengthen students' knowledge, insight, and skills (Marcotte & Gruppen, 2022). The integrative curriculum in question is a combination of the curriculum prepared by the national government with the international curriculum used or developed in schools. The existence of an integrative curriculum can support the implementation of bilingual programs. Bilingual means the use of two languages to interact with humans (Dos Santos, 2019; Tribushinina & Mackaaij, 2023). Implementing a bilingual program is one form of

realizing quality educational services to face global competition. Bilingual programs have been applied to various levels of education with the aim of improving language skills (Dos Santos, 2019; Molina, Salguero, Zhong, & Soosai Raj, 2023; Tribushinina & Mackaaij, 2023). Bilingual programs can direct students to improve their linguistic abilities and think and act creatively (Yemez & Dikilitaº, 2022). Bilingual programs in schools provide an optimal and intensive environment for second languages (Baron, Connell, & Griffin, 2022; Gillet, Barbu, & Poncelet, 2021).

In Indonesia, English has become a foreign language that must be studied. So, English can be used as a language of instruction in educational units to support students' foreign language skills. Junior high school is a critical stage in the formation of second language skills. Implementing a bilingual program at this level of education can help identify student needs and challenges. Bilingual in question is a learning process by integrating a foreign language and Indonesian. So, it can help students continue to a higher level of education. The implementation of bilingualism in junior high schools is carried out through various programs, both classroom learning programs and additional programs such as English Day. Bilingual programs are useful in improving students' abilities in mastering foreign languages.

One of the provinces in Indonesia that has innovated to implement a bilingual program is D.I. Yogyakarta, one of which is implemented at the private junior high school level. This is because the policy regarding the implementation of bilingual programs in state junior high schools has been abolished. Implementation of a bilingual program in mathematics and science subjects at private junior high school D.I. Yogyakarta is in line with the factors that encourage the importance of bilingual classes. This is because human resources are very necessary in creating competitiveness who master technology and basic knowledge, namely mathematics, and science and most of this knowledge is disseminated using English. Rationally, there are reasons that implementing bilingual classes can improve effective learning, especially in English, mathematics, and science competencies. However, the implementation of this bilingual program has not gone well. This is because it has not been implemented optimally by mathematics, science, and English teachers. Apart from that, school facilities also influence the implementation of bilingual programs, for example, the availability of a language laboratory. This is a challenge for schools that implement bilingual programs to improve the quality and quality of education so that they become superior schools as educational institutions.

Developing a process certainly requires an evaluation process because evaluation can measure students' capacity for bilingual mastery. Program evaluation must be carried out to determine the quality of a program (Irene, 2023). The evaluation of this learning program adopts the CIPP evaluation model. Specifically, evaluation in the context component is related to the objectives of implementing bilingual learning which includes English as the target language. Evaluation in the input component refers to the readiness of students, teachers, and facilities and infrastructure. Evaluation of the process components including the process of implementing learning in the school environment. Finally, evaluation in the product component refers to learning outcomes in mathematics and science subjects in bilingual classes.

METHODS

Participants

This evaluation was conducted in a private school because private schools have a flexible curriculum. This means that private schools can

design their own curriculum so that the implementation of the bilingual program can be adapted to the needs of the school and the wishes of the parents. In addition, some private schools have larger funds, so the facilities and resources provided are more supportive of the bilingual program. As in public schools, the regulations and education standards must follow the curriculum and education standards set by the local or national government. In addition, currently the bilingual program in public junior high schools has been abolished, so there are no more public junior high schools that implement bilingual programs. The population in this study were class VIII students of private junior high schools in D.I. Yogyakarta which have implemented a bilingual program including Budi Mulia Dua Middle School, Kesatuan Bangsa Middle School, Global Islamic School Middle School, Al-Azhar26 Islamic Middle School, LHI Banguntapan Integrated Islamic Middle School, Kesatuan Bangsa Bilingual Boarding School Middle School, Mutiara Persada Middle School, Tumbuh Middle School, and Muhammadiyah Middle School. 3 Yogyakarta. The sampling technique from the population used purposive sampling with a total of 52 students as respondents. Apart from that, there were bilingual class teachers and those in charge of the curriculum as respondents. Bilingual class teacher to provide detailed information and objective assessment of program implementation. The person in charge of the curriculum is the party who provides verification of the bilingual program implementation data.

Research Design and Procedures

This research is descriptive quantitative research. Descriptive quantitative aims to describe a situation objectively by using numbers to interpret the data obtained. This program evaluation model uses the CIPP evaluation model

The CIPP model is used to evaluate educational programs which consist of context, input, process, and product components (Ahmadipour, Alirezaie, & Mobasher, 2023; Irene, 2023; Meiklejohn dkk., 2023; Rebia, Suharno, Tamrin, & Akhyar, 2023). The context component plans activities before they start and becomes the basis for implementing the activity, the input component is the plan or strategy used to implement the program and aims to determine whether resources are sufficient to achieve program goals, the process component, the method used by the program to implement the progress of a program, and the product component. final results of program evaluation (Duan, Xiang, Niu, & Han, 2023; Fan, Tian, Lu, & Cao, 2022; Irene, 2023; Rama dkk., 2023; Suryadin, Nurfitriani, & Sari, 2022). The CIPP model in this research is as follows:

Context : *Includes school profile, foundation and objectives for implementing bilingual learning.*

Input : *Includes human resources such as language teachers and students as well as facility readiness and infrastructure.*

Process : *Includes language interaction, student involvement in learning, application of models and bilingual learning methods, learning barriers, and supporting activities language program.*

Product: Includes learning outcomes in terms of mathematics and science test results for both bilingual and national (regular) classes.

Instrument

In this research, data was collected using questionnaires, observations, and interviews. A questionnaire is a data collection technique by providing written statements or questions for respondents to answer (Klement & Terlau, 2022; Knickenberg, Zurbriggen, & Schwab, 2022;

Liantori & Sujarwo, 2023). The questionnaire used is a closed questionnaire looking for the percentage of questionnaire results (Sugiyono, 2019). Observations are used during the learning process to obtain information regarding the suitability of the teaching modules of the bilingual program to the actual situation in the classroom. Interview guide, this data collection instrument is a data source that supports questionnaires that are responded to by students and teachers and through interview guides can reveal phenomena from the components of context, input, process, and product. The instrument is a self-developed non-test instrument. There are 30 questions for teacher observation with indicators of introductory activities, core activities, and closing. In the learner questionnaire instrument, there are 20 items with indicators of resource readiness, language interaction, learning methods, and supporting factors. The learner questionnaire has 15 items with indicators of resource readiness, language interaction, learning models and methods, supporting and inhibiting factors, and program results in both bilingual and regular classes. In the curriculum interview guide, there are 10 items with indicators of bilingual program objectives, resource readiness, supporting factors, and program results in both bilingual and regular classes. In the bilingual teacher interview guidelines, there are 18 items with indicators of bilingual program objectives, resource readiness, language interaction, learning models and methods, supporting and inhibiting factors, and program results in both bilingual and regular classes.

Data Analysis

This research uses data analysis in the form of content validity, construct validity, reliability, and percentage of success. Validity is the accuracy of a test in carrying out its measurement function (Rahmawati, 2023). Content validity analysis was carried out by expert judgment and the results were processed using the Aiken formula (Astuti & Setiawan, 2023; Sarkity, Fernando, & Hindrasti, 2023). The Validity Index categorizes items into three levels based on their validity scores: High (if >0.8), indicating strong validity; Medium (0.4 – 0.8), representing moderate valid; and Low (if >0.4), suggesting lower validity.

Construct validity uses exploratory factor analysis to confirm whether the factors formed are supported by data and can prove that the statement items do measure the specified factors (Astuti & Setiawan, 2023; Rahmawati, 2023; Syaifuddin, Sarwi, Hartono, & Nuswowati, 2023). Reliability is used to see how consistent the measurement results are. Reliability in this calculation uses Cronbach's Alpha formula for the reason of using a Likert scale questionnaire (Retnawati, 2016). This test uses SPSS 16.10 with Cronbach's Alpha technique. The decision criteria in Cronbach's Alpha formula range from 0 to 1, with a cut-off value of 0.7 (Fauzan, Triyono, Hardiyanta, Daryono, & Arifah, 2023; Hariyanto, Daryono, Hidayat, Hadi Prayitno, & Nurtanto, 2022; Widyastuti, Hadi, Daryono, & Abd Samad, 2023). Interview guide instruments, observation sheets, and questionnaires Reliability estimates were calculated using the Intraclass formula Correlation Coefficient (ICC). ICC shows the comparison between variations which results in the attribute being measured with variations in measurement whole. Meanwhile, the reliability criteria are guidance instruments interviews, observation sheets, and questionnaires with ICC reliability categories using the Fleiss Kappa scale (Widhiarso, 2011).

The program success criteria are categorized based on the percentage (Arikunto & Jabar, 2018). A range of 80 – 100% falls into the "Very good" category, indicating a highly successful program. A percentage between 66 and 79% is classified as "Good," signifying a program that has achieved a commendable level of success. The range of 56 – 65% is labeled as

Table 2. Fleiss kappa scale coefficient

Reliability coefficient	Category
K < 0.00	Poor agreement
0.00 < K < 0.20	Slight
0.21 < K < 0.40	Fair
0.41 < K < 0.60	Moderate
0.61 < K < 0.80	Substantial
0.81 < K< 1.00	Almost perfect agreement

"Quite good," suggesting a program with satisfactory success. A percentage between 40 and 55% is categorized as "Unfavorable," indicating a less successful program. Percentages below 40% are categorized as "Very poor," representing a program with significant room for improvement.

■ RESULTS AND DISCUSSION

Content Validity

Content validity analysis was conducted through expert judgment, and the results were processed using the Aiken formula. The Validity Index categorizes items into three levels based on their validity scores: High (V > 0.8), indicating high validity; Medium (0.4 - 0.8), reflecting moderate validity; and Low (V > 0.4), indicating lower validity. The validation results for the Observation instrument showed a CVI range of 0.75 - 1.00, categorizing it as medium-high. Expert assessment of the 30 observation instrument items declared it comprehensively valid, with 2 items in the medium category and 28 items in the high category. For the Questionnaire Content Validation, all 15 items had a CVI value > 0.80, indicating the content validity of the 15-item questionnaire and placing it in the high category. The Field Interview instrument, consisting of 10 items, demonstrated a CVI range of 0.75 - 1.00, categorizing it as medium-high. Nine items had values above 0.8 and were included in the high category, while one item with a value of 0.75 fell into the medium category. The Bilingual Teacher Interview, with 18 instrument items, showed a CVI > 0.80, categorizing all items as high. Content validity for the questionnaire, comprising 20 instrument items, was confirmed with all items declared valid in the high category, as each item's value exceeded 0.8.

Construct Validity

The results of the content validity analysis were conducted through expert judgment, and the results were processed with Aiken's formula. The Validity Index categorizes items into three levels based on validity scores: High (CVI > 0.8), indicating high validity; Medium (0.4 - 0.8), reflecting moderate validity; and Low (CVI > 0.4), signifying lower validity. The validation results for the Observation instrument showed a CVI range of 0.75 - 1.00, which was categorized as medium-high. Expert assessment of the 30 items of the observation instrument stated that it was comprehensively valid, with 2 items in the medium category and 28 items in the high category.

The Anti-Image Correlation Coefficient of the Construct Validity Questionnaire obtained MSA values ranging from 0 to 1. If the MSA value = 1, then this variable can be predicted without error by other variables. If the MSA value is greater than 0.5, then the factors can still be analyzed. Meanwhile, if the MSA value is less than 0.5 and close to zero, then the variable cannot be analyzed or excluded. From the analysis results, there are two items with values below 0.5, namely item number 2 (0.472) and item number 17 (0.351). However, these two items were not excluded because this study used a pilot test.

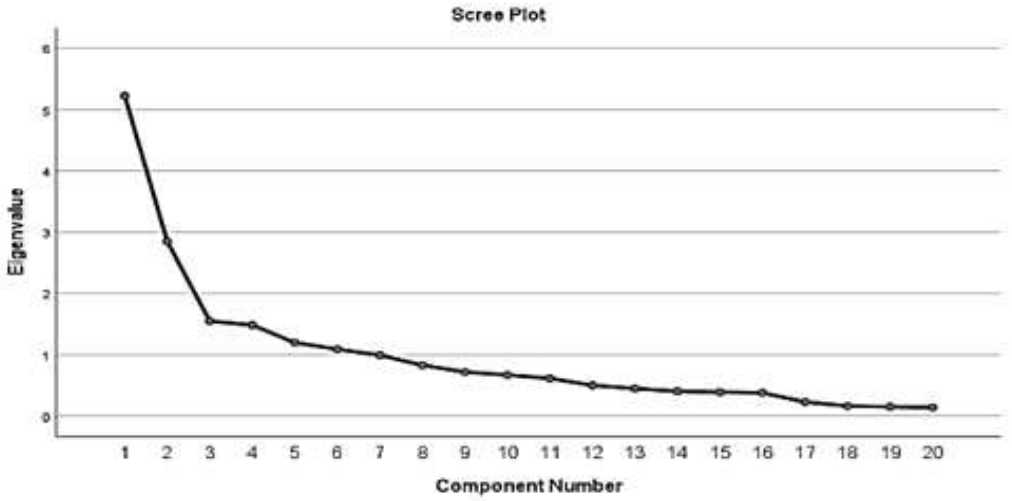

Figure 1. Scree plot of questionnaire construct validity

Based on the eigenvalues and variance components of the analysis results factors using SPSS can be obtained from questionnaire data students towards bilingual learning contains 6 more eigenvalues greater than 1, so it can be said that bilingual learning contains 6 (six) factors. Of these six factors, there is 66.945% of the variance which can be explained. However, by paying attention to the Scree Plot of eigenvalues, a graph consisting of only 2 (two) slopes is obtained, while others show a sloping graph. These results indicate that there are 2 (two) dominant factors that were measured in this instrument. The dominant factors are input and process.

Table 1. Rotated component matrix coefficients

Component Matrix[a]						
	Component					
Items	**1**	**2**	**3**	**4**	**5**	**6**
1			0.833			
2				0.473		
3	0.492	0.512				
4			0.811			
5	0.468	0.529				
6				0.844		
7				0.673		
8					0.681	
9			0.529			
10	0.652					-0/406
11		0.692				
12	0.783					
13		0.784				
14					0.584	
15	0.627					
16			0.687			
17						0.820

18	0.522			0.543
19		0.725		
20				0.605

The rotated Component Matrix contains a factor loading value > 0.4 on the main factor, a factor loading value on alternative factors < 0.3, and there is a difference of at least 0.2 between the main factor and alternative factors. This is useful for knowing which items are classified by the factors formed. Based on the Rotated Component Matrix it is known that there are 6 factors that influence the 20 questionnaire items. On factor 1 obtained items 10, 12, and 15. Followed by items 3, 5, 11, 13, and 19 are included in factor 2. Followed by items 1, 4, 9, and 16 which entered into factor 3. As for factor 4, items 2, 6, and 7 were obtained. Next, factor 5 consists of items 8, 14, 18, and 20. Meanwhile, factor 6 consists of item 17.

Reliability

The results of the calculation of the estimated reliability of the questionnaire instrument in this study show that the questionnaire has a level of reliability in the Almost Perfect Agreement category, namely 0.812. Therefore, questionnaires can be used to collect research data. Next is the calculation Reliability estimation on the interview guide instrument using inter-rater reliability which shows the results that the interview guide instrument (curriculum area) 0.838 and the interview guide (teacher) 0.784 have high reliability, the reliability of the observation instrument level medium reliability, namely 0.715 and high reliability of the questionnaire instrument, namely 0.767.

Table 2. Instrument reliability estimation results

Instrument	Coefficient Reliability	Information
Questionnaire	0.812	Reliable
Interview guide (teacher)	0.784	Reliable
Interview guide (teacher bilingual)	0.838	Reliable
Questionnaire	0.767	Reliable
Observation sheet	0.715	Reliable

Evaluation of Bilingual Programs with the CIPP Model

Context

Evaluation of the context component of Budi Mulia Dua Middle School based on documentation review in terms of the school profile, Budi Mulia Dua Middle School has a vision and mission that supports the objectives of implementing bilingual learning. Students' universal values and knowledge can be achieved through a good command of English. Thus, the vision and mission of Budi Mulia Dua Middle School is towards bilingual learning. The foundation of Budi Mulia Dua Middle School is based on the school's vision and mission to make students become world citizens. The aim of implementing the bilingual program is to improve the student learning process. The foundation and objectives are also seen from the Budi Mulia Dua Middle School process in planning and compiling the learning curriculum. In the curriculum planning process, teachers dissect the national and Cambridge curricula to see the suitability of learning materials to student needs.

Evaluation of the context component of Mutiara Persada Middle School based on a

review of school profile documentation, Mutiara Persada Middle School has a vision and curriculum that supports the implementation of bilingual learning. The bilingual learning program has goals and foundations that lead to bilingual learning to produce students who excel at national and international levels, requiring a good command of the English language. The foundation and aim of holding a bilingual learning program at Mutiara Persada Middle School is to seeing the opportunity that English is an international language so students must have high English language skills.

Evaluation of the SMP Tumbuh context component is based on a documentation review of the school's profile, which has a vision of children growing and developing as learners who have inclusive character, care about the environment, respect local wisdom, have a national perspective, and are globally competitive. The aim of SMP Tumbuh is to have educational goals that lead to bilingual learning. This is because producing students who are globally competitive requires a good command of English.

Evaluation of the context of implementing the bilingual learning program at Budi Mulia Dua Middle School, Mutiara Persada Middle School, and Tumbuh Middle School is measured by indicators of the school profile, the objectives of implementing the bilingual program, and the basis for holding the bilingual program. The evaluation results show that the three schools' learning programs have formulated a bilingual program based on these 3 indicators.

Input

Based on table 14, shows that the facilities and infrastructure of Budi Mulia Dua Middle School have very good success criteria with a percentage of 83%. This percentage shows that Budi Mulia Dua Middle School has adequate facilities and infrastructure. Then, regarding human resource readiness, the percentage of human resource readiness was 64%, which means that Budi Mulia Dua Middle School has quite good human resources. In this case, the teacher has the ability to operate a computer and is able to speak English well.

Table 3. Percentage of success of input components based on questionnaire results

No.	School Name	Input Indicator	
		Facilities and infrastructure (%)	HR Readiness (%)
1	SMP Budi Mulia Dua	83%	64%
2	SMP Mutiara Persada	68%	56%
3	SMP Tumbuh	84%	71%

The facilities and infrastructure at Mutiara Persada Middle School achieved a program success percentage of 68% and were in a good category. This means that Mutiara Persada Middle School has facilities and infrastructure that support bilingual learning. Meanwhile, the indicator of human resource readiness, namely teachers, at Mutiara Persada Middle School has met the criteria for a successful program in the quite good category. This can be seen from the percentage of human resource readiness which is 56% and is in the sufficient category.

Based on the percentage obtained, SMP Tumbuh has a program success of 84% in terms of facilities and infrastructure indicators. This percentage shows that the facilities and infrastructure owned by Tumbuh Middle School are in the very good category. Thus, Tumbuh

Middle School has very adequate facilities and infrastructure to support bilingual learning.

The availability of facilities and infrastructure for implementing the bilingual learning program through questionnaires and interviews shows a quite satisfactory level of achievement with presentation results with good success criteria. The interview results show that the school has provided various facilities such as projectors, wifi, bilingual books, PCs, and language labs to support bilingual learning at school. The readiness of human resources in schools has met the criteria for satisfactory achievement with a good percentage and is supported by qualitative data. This can be seen from the teacher's skills in using technology and the ability to explain learning material using English fluently. Thus, 3 schools have prepared competent human resources in implementing bilingual learning.

Process

Based on table 15 from SMP Budi Mulia Dua, it shows that the indicators of teacher interaction with students and fellow teachers get a percentage of 72% in the good category, student participation gets a percentage of 57% in the sufficient category, learning methods get a percentage of 80% in the excellent category, and supporting factors for the implementation of bilingual programs have a percentage of 70% in the good category. Based on the percentage obtained, Mutiara Persada Junior High School has a component of process success in teacher interaction with students and fellow teachers of 76% in the good category, student participation of 75% in the good category, learning methods of 77% in the good category, and supporting factors for the implementation of the bilingual program of 82% and is in the excellent category. Furthermore, the percentage obtained at Tumbuh Junior High School shows that teacher interactions with fellow students have a percentage of 83% in the excellent category, student participation has a percentage of 67% in the sufficient category, learning methods, and supporting factors for the implementation of bilingual programs have a percentage of 83% in the excellent category.

Table 4. Percentage of process component success

No.	School Name	Process Indicator			
		Teacher Interaction with Students and Each Other (%)	Student Participation (%)	Learning methods (%)	Supporting Factors for Bilingual Programs (%)
1	SMP Budi Mulia Dua	72%	57%	80%	70%
2	SMP Mutiara Persada	76%	75%	77%	82%
3	SMP Tumbuh	83%	67%	83%	83%

The evaluation reviewed from 3 schools in the process component had teacher interaction with students and each other, student participation, bilingual models and learning, as well as supporting and inhibiting factors for implementing the bilingual program. This indicates that these schools have implemented bilingual learning programs well. The use of language in this bilingual class is dominated by the use of English. The learning carried out does not only focus on teachers and students, but there is the interaction between teachers and students. This interaction occurs during teaching and learning activities. Student participation has a positive impact because it creates dynamic and competitive learning. In these three schools, learning methods and models have been implemented that can support the implementation

of bilingual learning programs. So, effective methods and models can help students learn English lesson concepts. Apart from that, in implementing a bilingual learning program, various supporting and inhibiting factors are determined, such as the habit of speaking in public using English, the availability of facilities and infrastructure, and the availability of training for educators.

Product

Based on Table 16, the assessment of the Budi Mulia Dua Middle School product components is seen based on the PAT results for

Table 5. Average achievement of learning outcomes in mathematics and science subjects

No.	School Name	Mathematics		IPA	
		Bilingual	National	Bilingual	National
1	SMP Budi Mulia Dua	80.07	77.78	82.51	77.65
2	SMP Mutiara Persada	79.5	78.36	83.3	81.26
3	SMP Tumbuh	84.2	75.15	80.6	77.38

mathematics and science subjects in bilingual and national classes, obtaining an average result of 80.07 in the bilingual class and an average score of 77.78 in the national class in mathematics subjects. Meanwhile, in science subjects, the bilingual class got a score of 82.51 and the national class got a score of 77.65.

Then at Mutiara Persada Middle School, PAT results were obtained for mathematics subjects in the bilingual class with an average score of 79.5 and in the national class with an average score of 78.36. Meanwhile, in science subjects the bilingual class obtained an average result of 83.3 and the national class obtained an average result of 81.26.

The average achievement of junior high school learning outcomes is Tumbuh, in mathematics subjects the bilingual class gets an average score of 84.2 and the national class gets an average score of 75.15. Meanwhile, in science subjects the bilingual class obtained an average result of 80.6 and the national class obtained an average result of 77.38.

Judging from the product component, learning results from PAT scores for the 2022/2023 academic year show that students in bilingual classes have higher average learning outcomes than national classes. This shows that the product resulting from the bilingual program is said to be successful.

■ CONCLUSIONS

Based on the results of the evaluation of the bilingual learning program at Budi Mulia Dua Middle School, Mutiara Persada Middle School, and Tumbuh Middle School, it was found that the context evaluation of the bilingual program was in accordance with the school's vision, mission, goals and curriculum, the input evaluation for the bilingual program obtained an average percentage of below 100% because the facilities and infrastructure and human resources available are not optimal, the evaluation process in the bilingual program has not reached 100% because the English language skills of students and teachers are not yet optimal, and the evaluation of the product component of bilingual class student learning outcomes is higher than national class. This can be seen from the average PAT scores for mathematics and science subjects in bilingual classes and national classes.

■ REFERENCES

Ahmadipour, H., Alirezaie, A., & Mobasher, M. (2023). Psychometric properties of the context, input, process, and product (cipp)

model for the evaluation of medical ethics education. *Strides in Development of Medical Education*, *20*(1), 155–161. https://doi.org/10.22062/sdme.2023.198729.1240

Arikunto, S., & Jabar, C. S. (2018). *Evaluasi program pendidikan.* Jakarta: PT Bumi Aksara.

Astuti, Y. A. Y., & Setiawan, A. (2023). *Pengembangan instrumen penilaian sikap percaya diri dalam pembelajaran bagi siswa sekolah dasar. Tuladha: Jurnal Pendidikan Dasar*, *2*(2), 122–138. https://doi.org/10.30738/tuladha.v2i2.15976

Baron, A., Connell, K., & Griffin, Z. M. (2022). Grammatical gender in spoken word recognition in school-age spanish-english bilingual children. *Frontiers in Psychology*, *13*(February), 1–10. https://doi.org/10.3389/fpsyg.2022.788076

Dos Santos, L. M. (2019). Bilingual English education: Expectation of parents who enrol their children in Bilingual primary schools. *International Journal of Instruction*, *12*(4), 747–766. https://doi.org/10.29333/iji.2019.12448a

Duan, P., Xiang, J., Niu, H., & Han, C. (2023). Construction of evaluation index for chinese engineering undergraduates based on cipp model. *SAGE Open*, *13*(1), 21582440221149416. https://doi.org/10.1177/21582440221149415

Fan, X., Tian, S., Lu, Z., & Cao, Y. (2022). Quality evaluation of entrepreneurship education in higher education based on CIPP model and AHP-FCE methods. *Frontiers in Psychology*, *13*(September), 1–12. https://doi.org/10.3389/fpsyg.2022.973511

Fauzan, A., Triyono, M. B., Hardiyanta, R. A. P., Daryono, R. W., & Arifah, S. (2023). The effect of internship and work motivation on students' work readiness in vocational education: PLS-SEM Approach. *Journal of Innovation in Educational and Cultural Research*, *4*(1). https://doi.org/10.46843/jiecr.v4i1.413

Gillet, S., Barbu, C., & Poncelet, M. (2021). Early bilingual immersion school program and cognitive development in French-speaking children: Effect of the second language learned (English vs. Dutch) and exposition duration (2 vs. 5 years). *PLOS ONE*, *16*(10), e0258458. Diambil dari https://doi.org/10.1371/journal.pone.0258458

Hariyanto, V. L., Daryono, R. W., Hidayat, N., Hadi Prayitno, S., & Nurtanto, M. (2022). A framework for measuring the level of achievement of vocational students competency of architecture education. *Journal of Technology and Science Education*, *12*(1). https://doi.org/10.3926/jotse.1188

Irene, E. A. (2023). Evaluation of teacher education curricula and its relevance to licensure examination using Context, Input, Process and Product (CIPP) model. *Social Sciences & Humanities Open*, *8*(1), 100607. https://doi.org/10.1016/j.ssaho.2023.100607

Klement, J., & Terlau, W. (2022). Education for sustainable development and meaningfulness: evidence from the questionnaire of eudaimonic well-being from german students. *International Journal of Environmental Research and Public Health*, Vol. 19. https://doi.org/10.3390/ijerph19116755

Knickenberg, M., Zurbriggen, C. L. A., & Schwab, S. (2022). Validation of the student version of the perceptions of inclusion questionnaire in primary and secondary education settings. *SAGE Open*, *12*(1), 21582440221079896. ttps:/

/doi.org/10.1177/21582440221079896

Liantori, B., & Sujarwo, S. (2023). Citizenship education in elementary schools: factors influencing classroom participations of college students. *Jurnal Pendidikan Progresif*, *13*(3), 1410–1421. https://doi.org/10.23960/jpp.v13.i3.202338

Marcotte, K. M., & Gruppen, L. D. (2022). Competency-based education as curriculum and assessment for integrative learning. *Education Sciences*, *12*(4). https://doi.org/10.3390/educsci12040267

Meiklejohn, S., Anderson, A., Brock, T., Kumar, A., Maddock, B., Wright, C., … Kent, F. (2023). The utility of an interprofessional education framework and its impacts upon perceived readiness of graduates for collaborative practice. A multimethod evaluation using the context, input, process, product (CIPP) model. *Nurse Education Today*, *121*, 105707. https://doi.org/10.1016/j.nedt.2023.105707

Molina, I. V., Salguero, A., Zhong, S., & Soosai Raj, A. G. (2023). The effects of spanish-english bilingual instruction in a cs0 course for high school students. *Proceedings of the 2023 Conference on Innovation and Technology in Computer Science Education V. 1*, 75–81. Diambil dari https://doi.org/10.11453587102.3588845

Rahmawati, E. (2023). *Pengembangan instrumen sikap nasionalis pada peserta didik sekolah dasar* [development of nationalist attitude instruments for elementary school learners]. *PUSTAKA: Jurnal Bahasa dan Pendidikan*, *3*(1), 171–183. Diambil dari https://jurnal-stiepari.ac.id/index.php/pustaka/article/view/830

Rama, A., Ambiyar, A., Rizal, F., Jalinus, N., Waskito, W., & Wulansari, R. E. (2023). *Konsep model evaluasi context, input, process dan product (CIPP) di sekolah menengah kejuruan* [The concept of context, input, process and product (CIPP) evaluation model in vocational high schools]. *JRTI (Jurnal Riset Tindakan Indonesia)*, *8*(1), 82. https://doi.org/10.29210/30032976000

Rebia, P. S., Suharno, Tamrin, A. G., & Akhyar, M. (2023). Evaluation of product-based education training class at vocational high school using the cipp model. *Journal of Curriculum and Teaching*, *12*(3), 135–146. https://doi.org/10.543/jct.v12n3p135

Retnawati, H. (2016). *Analisis kuantitatif instrumen penelitian*. Yogyakarta: Parama Publishing.

Sarkity, D., Fernando, A., & Hindrasti, N. E. K. (2023). Designing and Content Validity of Instrument for Measuring Marine Environmental Care Attitude Through Integrated Science Learning. *BIO Web Conf.*, *70*. https://doi.org/10.1051/bioconf/20237002011

Sugiyono. (2019). *Metode penelitian kuantitatif kualitatif dan R&D*. Bandung: ALFABETA.

Suryadin, A., Nurfitriani, N., & Sari, W. P. (2022). Lazismu and muhammadiyah philanthropy in the covid-19 pandemic (CIPP Evaluation Model). *Jurnal Ilmiah Peuradeun; Vol 10 No 2 (2022): Jurnal Ilmiah Peuradeun*. https://doi.org/10.26811/peuradeun.v10i2.674

Syaifuddin, S., Sarwi, S., Hartono, H., & Nuswowati, M. (2023). Karakterisasi Instrumen Tes Keterampilan Proses Sains pada Bahan Kajian Fluida Dinamis Berdasarkan Teori Tes Klasik [Characterization of Science Process Skills Test Instrument on Dynamic Fluid Study Material Based on Classical Test Theory]. *Prosiding Seminar Nasional Pascasarjana*, *6*(1), 818–825.

Tribushinina, E., & Mackaaij, M. (2023).

Bilingual advantages in foreign language learning: evidence from primary-school pupils with developmental language disorder. *Frontiers in Education*, Vol. 8. https://doi.org/10.3389/feduc.2023.1264120

Tulomana, J., Tagimaucia, V., & Chand, S. P. (2023). Assessing the readiness of pre-service teachers to teach english to esl learners in fijian secondary schools: a study of their language learning experiences and perceptions. *BELT - Brazilian English Language Teaching Journal*, *14*(1 SE-Article), e44931. https://doi.org/10.15448/2178-3640.2023.1.44931

Uccelli, P. (2023). Midadolescents' language learning at school: toward more just and scientifically rigorous practices in research and education. *Language Learning*, *73*(S2), 182–221. https://doi.org/10.1111/lang.12558

Widhiarso, W. (2011). *Melibatkan rater dalam pengembangan alat ukur*. Yogyakarta: Universitas Gadjah Mada Press.

Widyastuti, P., Hadi, S., Daryono, R. W., & Abd Samad, N. Bin. (2023). The mediation role of university environment in the relationship between self-efficacy and family environment on entrepreneurial education interest: A PLS-SEM Approach. *Indonesian Journal on Learning and Advanced Education (IJOLAE)*, *5*(3), 295–310. https://doi.org/10.23917/ijolae.v5i3.22015

Yemez, N., & Dikilitaº, K. (2022). Development of verbal creativity by bilingual and english as foreign language learners in kindergarten to 8th grade schools. *Creativity Studies*, *15*(1), 25–39. https://doi.org/10.3846/cs.2022.12603