Research Article

Development of Digital-based Program Evaluation Model (McIPP): A Study on Elementary School Indonesian Language Learning

Anlianna[1], Erman[2*], Hendratno[3]

[1,2,3]Universitas Negeri Surabaya, Indonesia

Corresponding Author: Erman, Email: erman@unesa.ac.id

[1]anlianna.22001@mhs.unesa.ac.id

[2]hendratno@unesa.ac.id

[3]erman@unesa.ac.id

ARTICLE INFO	ABSTRACT
	Introduction: Traditional models of program evaluations are less responsive to learning in the 21st century, particularly with technology incorporation and precision in data. Hence, the construction of the digital-based evaluation model (McIPP) plays an important role in the improvement of Indonesian language skill learning at the Elementary School level. **Objectives**: The goal of this study was to evaluate the content validity of the digital McIPP Evaluation Model through the Aiken's V method. Additionally, the study aimed to assess the construct validity of the model using Confirmatory Factor Analysis (CFA). **Methods**: This research is a type of development study aimed at creating a valid and practical digital-based McIPP program evaluation model for use in Indonesian language instruction at the elementary school level. The study follows an adapted version of the Borg & Gall development framework, consisting of four primary phases: a literature review and needs assessment, initial model and indicator design, expert evaluation, and empirical validation of the model's constructs. Participants in this research included experts in educational evaluation, instructional technology, and Indonesian language for content validation, along with principals, teachers, and students for construct validation. The study involved 15 school principals, 86 teachers, and 508 students from 15 elementary schools representing high, medium, and low quality categories in Pasuruan City and Southeast Sulawesi Province. Data were gathered through the McIPP evaluation questionnaire. Content validity was evaluated using Aiken's V, with a threshold of 0.70 or above, while construct validity was assessed through Confirmatory Factor Analysis (CFA), utilizing standard loading factor criteria to evaluate the model's fit. **Results**: The instrument used in this study demonstrates strong content validity, as indicated by Aiken's V scores ≥ 0.08 across all items, with values ranging between 0 and 1.00 confirming the validity of the content in the digital-based evaluation model. Additionally, results from the Confirmatory Factor Analysis (CFA) indicate that the McIPP digital evaluation model aligns well with the empirical data. This is supported by model fit indices that meet established criteria, including Chi-Square/df (1.381), RMSEA (0.058), SRMR (0.072), and other key indicators such as NFI, CFI, TLI, and GFI, all of which score above 0.90. **Conclusions**: In conclusion, this study confirms that the digital-based McIPP evaluation model satisfies the validity criteria for both content and construct. As a result, it is considered suitable for implementation in the context of contemporary, responsive digital education evaluation. **Keywords:** digital based program evaluation model McIPP, Aiken's V, CFA, Indonesian Language

Research Article

INTRODUCTION

In the current digital landscape, the success of educational initiatives largely hinges on how effectively learning approaches can be reviewed and adapted in a timely manner [1], [2], [3]. At the elementary school level, learning Indonesian plays a vital role in building students' foundational literacy skills, which are essential for their success in other academic areas [4]. Despite this, many schools continue to rely on traditional evaluation methods that do not keep pace with technological advancements and are limited in providing the detailed data needed for program enhancement.

When schools continue to use traditional evaluation methods that fail to adapt to technological progress and struggle to deliver detailed data, a range of negative consequences may occur [5], [6], [7]. Firstly, educational decisions may become misguided, as subjective assessments, lacking real-time data, hinder the identification of areas that require improvement in teaching programs [8]. Secondly, progress in teaching is hindered, as educators and schools struggle to pinpoint areas for growth. This limits the implementation of differentiated learning approaches, personalized instruction, and the integration of technology [9]. Thirdly, traditional models fail to deliver quick and detailed feedback for both educators and students, leading to slow progress in teaching quality and students' mastery of skills. This is particularly evident in learning Indonesian, which demands critical thinking and active language abilities [10]. Fourth, the effectiveness of the Independent Curriculum is limited, as formative evaluations and diagnostic assessments cannot be fully implemented without sufficient technological and data support [11]. Finally, the disparity in school quality is growing, as institutions that do not embrace digital evaluation methods are falling behind in learning outcomes and teaching effectiveness compared to those that are more adaptable to technological advancements [6], [7], [12].

Evaluating learning programs through digital means is crucial for enhancing educational quality, as it provides precise, real-time data that is easy to analyze and supports informed decision-making [13], [14]. Leveraging digital technology makes the evaluation process more efficient, transparent, and inclusive, while also allowing for ongoing tracking of student progress [15]. Digital evaluation also fosters a more flexible and personalized learning experience, enabling educators to swiftly recognize students' strengths and weaknesses and modify their teaching methods accordingly [16]. Moreover, digital-based evaluation promotes the use of technology in education, speeds up innovation, and streamlines the documentation and reporting of learning outcomes [17]. In the implementation of the Independent Curriculum, digital evaluation serves as a crucial tool for objectively and systematically assessing the effectiveness of contextual and differentiated learning.

However, initial studies conducted in elementary schools under the Pasuruan City Education Office and the Education Offices of Southeast Sulawesi Province revealed confusion regarding the implementation of learning evaluations. There is a need for a simple, technology-based evaluation program. Discussions with school principals, teachers, and technology experts indicated that a development approach using Microsoft Forms, in combination with the CIPP (Context, Input, Process, and Product) evaluation model, was necessary. The research found that the CIPP model was effective in assessing the quality of teaching [18]. Evaluate Online Teacher Training Program with CIPP [19]. The CIPP evaluation model, created by Daniel L. Stufflebeam, is designed to enhance educational programs, including classroom instruction, by focusing on four key components: Context, Input, Process, and Product [20]. This approach is holistic, as it evaluates not only the final outcomes but also takes into account learning needs, resource availability, the execution of learning activities, and student performance [21]. Therefore, the CIPP model facilitates ongoing enhancement of learning quality, aligning with students' needs [22].

Previous studies have not explored the development of a digital-based CIPP model. Building on the findings of preliminary research and earlier studies, the researcher aims to create a program evaluation model using McIPP (Microsoft Forms, Input, Process, and Product) technology. This model is expected to offer a more dynamic, measurable, and real-time evaluation system. McIPP enables the active participation of various stakeholders (teachers, students, and principals) in a cohesive manner, providing quicker, data-driven feedback. Developing this model is crucial to addressing the challenges of implementing the Independent Curriculum, which emphasizes contextual, differentiated learning and focuses on student outcomes. By incorporating a digital approach, the evaluation of Indonesian Language programs in elementary schools can be conducted more objectively, transparently, and accountably, ultimately contributing to the enhancement of basic education quality in Indonesia.

Research Article

OBJECTIVES

This study aims to design a digital-based program evaluation model using the McIPP approach (Microsoft Forms, Input, Process, and Product) that is both valid and effective for assessing Indonesian language instruction at the elementary school level. More specifically, the objectives of this research are to:

1. To evaluate the content validity of the digital-based McIPP Evaluation Model using Aiken's V analysis, involving expert reviewers in educational evaluation, instructional technology, and Indonesian language education. This process is intended to confirm that each indicator within the model is relevant, clearly defined, and sufficiently represents the dimensions of program evaluation.

2. The construct validity of the proposed model was assessed using Confirmatory Factor Analysis (CFA), with the goal of evaluating how well the key elements—Microsoft Forms, Input, Process, and Product—along with their related indicators, correspond to the theoretical framework of the evaluation model. CFA was also used to confirm the model's alignment with the data and to assess the strength of the relationships between the constructs and their respective indicators.

METHODS

This study follows a research and development (R&D) approach with the goal of creating a valid, digital-based McIPP program evaluation model suitable for assessing Indonesian language learning in elementary schools. The R&D methodology is adapted from the ten-step development model proposed by Borg and Gall, which has been streamlined into several key phases: (1) conducting a literature review and needs analysis, (2) designing the initial model and formulating its indicators, (3) validating the model through expert judgment, and (4) empirically testing the model's construct.

The subjects in this study included experts for content validity assessment, representing three areas of specialization: educational evaluation, instructional technology, and Indonesian language education. Additionally, the research involved principals, teachers, and students as participants. The sampling technique used was purposive random sampling, based on school quality classifications—high, medium, and low. A total of 15 elementary schools were selected as trial sites, consisting of five high-quality schools, five of moderate quality, and five lower-quality schools. These schools were located across regions under the jurisdiction of the Pasuruan City Education Office and the Education Offices of regencies/cities in Southeast Sulawesi Province. Based on this sampling, the study involved 15 principals, 86 teachers, and 508 students.

The data collection tool used in this study is a questionnaire specifically designed based on the McIPP evaluation model. It includes a set of indicators aimed at assessing digital-based Indonesian language learning programs. To evaluate content validity, expert reviews were conducted, while the construct validity was examined by distributing the questionnaire to principals, teachers, and students

The data analysis in this research involves examining both content and construct validity. Content validity was evaluated using Aiken's V method, which calculates the relevance of each indicator based on expert judgments. An indicator is considered valid if the Aiken's V score is equal to or greater than 0.80 Aiken on [23]. Construct validity is assessed using the Confirmatory Factor Analysis (CFA) method, supported by statistical software such as AMOS. An indicator is considered valid when the model's fit indices align with established standard criteria [24].

RESULTS

The outcomes of developing the McIPP digital-based program evaluation model in this study encompass both content validity, assessed through Aiken's V, and construct validity, evaluated using Confirmatory Factor Analysis (CFA).

Tabel 1. Content Validity of the McIPP Digital-Based Program Evaluation Model

No	Rated aspect	Aiken's V Value	Content validity
1	Clarity of the contents of the general evaluation instructions	0,85	Valid
2	Clarity of the steps of the assessment process	0,90	Valid

Research Article

No	Rated aspect	Aiken's V Value	Content validity
3	Clarity of evaluation results recommendations	0.85	Valid
4	Clarity on the timing of the evaluation	0.80	Valid
5	Clarity on how scores are processed	0.85	Valid
6	Completeness of the contents of the guide	0.85	Valid
7	Effectiveness of standard sentences	0.90	Valid
8	Accurate use of spelling and punctuation	0.90	Valid
9	Use of easy to understand words/terms	0.80	Valid
10	Shape and size of letters	0.80	Valid

According to the range of Aiken's V values between 0 and 1.00, as shown in Table 4.17, the research instrument demonstrates strong content validity, with all Aiken's V values being ≥ 0.8. The results in Table 1 confirm that the Aiken's V values fall within the 0 to 1.00 range, validating the content of the digital-based program evaluation model.

Tabel 2. Goodness-of-Fit Indeks CFA Results

Indeks Fit	Value	Standard Criteria	Description
ChiSqr/df	1,381	< 3,00	Good
RMSEA (Root Mean Square Error of Approximation)	0,058	≤ 0,08	Good
NFI (Normed Fit Index)	0,915	> 0,90	Good
CFI (Comparative Fit Index)	0,974	> 0,90	Good
TLI (Tucker-Lewis Index)	0,954	> 0,90	Good
GFI (Goodness of Fit Index)	0,962	> 0,90	Good
SRMR	0,072	≤ 0,08	Good

The CFA analysis results presented in Table 2 show that the digital-based McIPP evaluation model fits the empirical data very well. All model fit indices meet the required criteria, indicating a good model fit. The Chi-Square/df value of 1.381 is below the maximum acceptable limit of 3.00, confirming a suitable model fit. The RMSEA value of 0.058 and SRMR of 0.072 are within the recommended tolerance (≤ 0.08), suggesting minimal model error. Additional fit indices, such as NFI (0.915), CFI (0.974), TLI (0.954), and GFI (0.962), all exceed the minimum threshold of 0.90, reinforcing the model's strong fit. Therefore, it can be concluded that the McIPP model demonstrates good construct validity and is appropriate for evaluating Indonesian language learning programs in elementary schools. Furthermore, it aligns well with empirical data, as evidenced by all fit indices meeting the established standards.

DISCUSSION

The content and construct validity assessments of the Digital-Based Program Evaluation Model (McIPP) demonstrate that this model is well-suited for evaluating Indonesian language learning programs in elementary schools. The content validity test, performed using Aiken's V, showed that all indicators had a V value of ≥ 0.80, signaling strong agreement among experts regarding the relevance, clarity, and adequacy of the indicators for the program's evaluation objectives. The high Aiken's V scores also suggest that the indicators effectively represent key aspects of organizing and implementing Indonesian language learning, in line with technological advancements and the requirements of the Independent Curriculum. In addition to content validity, the construct validity of the McIPP digital evaluation model was assessed using Confirmatory Factor Analysis (CFA), which revealed a strong fit with the empirical data, as all fit indices met the established criteria.

The findings of this study suggest that the development of a digital-based McIPP program evaluation model aligns well with the fundamental principles of comprehensive educational evaluation, addressing the learning needs of today. Using the CIPP (Context, Input, Process, Product) model for evaluation is crucial in education as it offers a holistic view of learning effectiveness, considering not just the final outcomes but also the processes and factors that support the learning experience [25], [26]. Effective learning is influenced by how well the process, planning, and

Research Article

learning environment facilitate students in achieving the desired competencies. The CIPP evaluation model is crucial in this context, as it offers a structured, holistic, and improvement-focused approach to assessment [27], [28]. The digital-based McIPP model integrates the efficient use of Microsoft Forms with a comprehensive evaluation framework comprising four key dimensions: Input, Process, and Output, which align with the needs of contemporary educational assessment. By leveraging Microsoft Forms, this model allows for real-time data collection, enhancing efficiency and accessibility. It enables teachers, students, and other stakeholders to complete evaluation forms remotely and at their convenience [29]. The gathered data can be processed instantly and automatically, minimizing the risk of human errors and accelerating the analysis to detect trends or issues in the learning process [30]. Moreover, the flexibility of Microsoft Forms enables the McIPP model to be implemented on a large scale without significant managerial constraints. In terms of content, the input dimension examines the readiness of resources, such as the curriculum, materials, teacher skills, and technological infrastructure, all of which are crucial in the digital learning era. The process dimension evaluates the effectiveness of teaching practices, the methods employed, and the integration of technology, while the output dimension focuses on student learning outcomes, particularly language skills and critical thinking in a digital setting. These three dimensions work together, aligning the McIPP model with the principles of objective, accountable, and data-driven evaluation aimed at improving quality. Therefore, the combination of digital tools and a comprehensive evaluative framework within the McIPP model makes it highly applicable to modern educational systems, especially in supporting adaptive and sustainable Indonesian language learning in elementary schools.

CONCLUSION

In conclusion, the digital-based McIPP evaluation model meets the validity standards in both its content and structure. It is not only theoretically sound but has also demonstrated empirical consistency in assessing the quality and effectiveness of Indonesian language learning programs in elementary schools. This model addresses the limitations of traditional evaluation methods, which often lack adaptability to technology and fail to provide precise, real-time data. As such, the McIPP model is a valuable tool for implementing a more modern and responsive approach to digital-based educational evaluation.

REFRENCES

[1] I. Muthi, "Evaluasi Program Digital Smart Classroom: Pendekatan CIPP Untuk Meningkatkan Kualitas Pembelajaran," *Jurnal Education and development*, vol. 12, no. 3, p. 399, 2024, doi: 10.37081/ed.v12i3.6073.

[2] G. Mohyuddin, M. A. Khan, A. Haseeb, S. Mahpara, M. Waseem, and A. M. Saleh, "Evaluation of Machine Learning Approaches for Precision Farming in Smart Agriculture System: A Comprehensive Review," *IEEE Access*, vol. 12, pp. 60155–60184, 2024, doi: 10.1109/ACCESS.2024.3390581.

[3] S. H. Mohammed *et al.*, "A Review on the Evaluation of Feature Selection Using Machine Learning for Cyber-Attack Detection in Smart Grid," *IEEE Access*, vol. 12, pp. 44023–44042, 2024, doi: 10.1109/ACCESS.2024.3370911.

[4] I. H. Rohmah and I. R. Julianto, "Strategi Guru Sekolah Dasar pada Pembelajaran Bahasa Indonesia Berbantuan Teknologi," *JURNAL CAHAYA EDUKASIA*, vol. 3, no. 1, pp. 49–53, 2025.

[5] A. T. Ananda, "Revitalisasi Pembelajaran PAI Melalui Teknologi Adaptif: Kajian Literatur Sistematis Era Society 5.0," *Edu Global: Jurnal Pendidikan Islam*, 2024, doi: 10.56874/eduglobal.v5i2.2064.

[6] D. Lee *et al.*, "The impact of generative AI on higher education learning and teaching: A study of educators' perspectives," *Computers and Education: Artificial Intelligence*, vol. 6, Jun. 2024, doi: 10.1016/j.caeai.2024.100221.

[7] M. A. Faheem, Z. Rayhan, A. Ullah, and R. Paudel, "AI-driven innovation In HRM and its impact on business management: An in depth study of technology advancement and strategic implementation," *Nanotechnol Percept*, vol. 20, no. S13, pp. 1174–1204, 2024, doi: 10.13140/RG.2.2.19108.49284.

[8] C. N. Miristianti, I. Sofiatin, and M. Iqbal, "Pengaruh Penggunaan Teknologi dalam Pembelajaran di Dunia Pendidikan Sekolah Dasar," *Insan Cendekia : Jurnal Pendidikan dan Kebudayaan*, vol. 1, no. 1, pp. 34–40, 2024, doi: 10.56480/insancendekia.v1i1_996.

[9] Luluk Baikuna *et al.*, "Peran Teknologi Pendidikan Dalam Pemanfaatan Pembelajaran IPS," *JURNAL PENDIDIKAN DAN ILMU SOSIAL (JUPENDIS)*, vol. 2, no. 1, pp. 102–115, Nov. 2023, doi: 10.54066/jupendis.v2i1.1182.

[10] Magvira and Nensilianti, "Pengaruh Penggunaan Media Berbasis Artificial Intelligence (AI) terhadap Kemampuan Berpikir Kritis dan Keterampilan Menulis Teks Eksposisi Siswa Kelas X UPT SMA Negeri 4 Luwu," *Jurnal Onoma: Pendidikan, Bahasa dan Sastra*, vol. 11, no. 2, pp. 1777–1799, 2025, [Online]. Available: https://e-journal.my.id/onoma

[11] H. Rustantono *et al.*, "Analisis Implementasi Kurikulum Merdeka Pada Pembelajaran IPS Di SMP PGRI 4 Tirtoyudo," *Jurnal Education and development*, vol. 12, no. 3, pp. 52–57, 2024, doi: 10.37081/ed.v12i3.6342.

[12] T. Depita, "Pemanfaatan Teknologi Dalam Pembelajaran Aktif (Active Learning) Untuk Meningkatkan Interaksi dan Keterlibatan Siswa," *TARQIYATUNA: Jurnal Pendidikan Agama Islam dan Madrasah Ibtidaiyah*, vol. 3, no. 1, pp. 55–64, Aug. 2024, doi: 10.36769/tarqiyatuna.v3i1.516.

[13] D. Amsari, S. S. Syam, F. Suciana, and E. Wahyuni, "College Students' Satisfaction on Digital-based Learning Evaluation Course in Elementary Education Program," *EDUMATSAINS*, vol. 9, no. 2, pp. 28–42, 2025, doi: 10.33541/edumatsains.

[14] S. Febriani, S. Zakir, D. Ilmi, and R. Setia Pramana, "Evaluasi Program Sekolah Digital dalam Meningkatkan Pemanfaatan Teknologi untuk Pembelajaran," *DIRASAH*, vol. 7, no. 2, 2024, [Online]. Available: https://ejournal.iaifa.ac.id/index.php/dirasah

[15] M. Azhar *et al.*, "Evaluasi Pembelajaran Bahasa Arab Berbasis Teknologi Digital," *Open Acces Ukazh: Journal of Arabic Studies*, vol. 6, no. 1, p. 2025, 2025, doi: 10.37274/ukazh.v6i1.1438.

[16] E. Abdi, I. Astuti, and Afandi, "Evaluasi Pembelajaran Matematika dalam Pemberdayaan Teknologi Digital Menggunakan Model Evaluasi CIPP di SMAN Pontianak," *Jurnal Alwatzikhoebillah : Kajian Islam, Pendidikan, Ekonomi, Humaniora*, vol. 11, no. 1, pp. 47–72, Jan. 2025, doi: 10.37567/alwatzikhoebillah.v11i1.2615.

[17] P. Nedungadi, R. Menon, G. Gutjahr, and R. Raman, "Evaluating Holistic Education and Digital Learning model for advancing SDG4: a longitudinal mixed-effects modeling approach," *Cogent Soc Sci*, vol. 10, no. 1, 2024, doi: 10.1080/23311886.2023.2299134.

[18] S. Zhang, L. Bai, and B. Yang, "Mixed teaching quality evaluation of organizational behavior course based on CIPP model," *Front Educ (Lausanne)*, vol. 10, 2025, doi: 10.3389/feduc.2025.1538539.

[19] K. M. Shaheen, "Using CIPP Model to Evaluate Online Teacher Training Programme for Single National Curriculum, 2021," *International Research Journal of Management and Social Sciences*, vol. V, no. 1, pp. 144–160, 2024.

[20] A. Ulfa Tenri Pada, M. Mawarpury, M. Saputri, and H. Yulisman, "Analysis of Students' Perception of the Implementation of E-Learning USK: An Evaluation Using the CIPP Model," in *KnE Social Sciences*, Knowledge E DMCC, Apr. 2024. doi: 10.18502/kss.v9i13.15994.

[21] G. Zhang *et al.*, "Using the Context, Input, Process, and Product Evaluation Model (CIPP) as a Comprehensive Framework to Guide the Planning, Implementation, and Assessment of Service-learning Programs Introduction and Review of Literature," 2011.

[22] S. Aziz, M. Mahmood, and Z. Rehman, "Article Journal of Education and Educational Developement Implementation of CIPP Model for Quality Evaluation at School Level: A Case Study CIPP Model for Quality Evaluation," *Journal of Education and Educational Developement*, vol. 5, no. 1, pp. 189–206, 2018.

[23] S. Nurjanah and W. Widihastuti, "The application of Aiken's V method for evaluating the content validity of instruments that measure the implementation of formative assessments," *Article in Journal of Educational Research and Evaluation*, 2023, doi: 10.15294/jere.v12i2.76451.

[24] D. Shahid Zeb, H. Said, and B. Bakri Badru, "Confirmatory Factor Analysis (Cfa) for Testing Validity And Reliability Instrument in the Study of Education," *Aust J Basic Appl Sci*, vol. 5, no. 12, pp. 1098–1103, 2011, [Online]. Available: https://www.researchgate.net/publication/261830668

[25] A. Sarfatra Winarno and J. Subando, "The Concept of Evaluating the Fiqh Learning Process Using the CIPP Model," 2025. [Online]. Available: https://jurnaldidaktika.org

Research Article

[26]   D. Larasati and M. Islami, "Evaluation of Learning Mathematics Economics and Business AE Class with Cipp Method," *Journal of Development Economic and Social Studies*, vol. 4, no. 1, pp. 52–60, Jan. 2025, doi: 10.21776/jdess.2025.04.1.05.

[27]   D. Suryana, A. Husna, and N. Mahyuddin, "CIPP Evaluation Model: Analysis of Education Implementation in PAUD Based on Government Policy on Implementation of Learning During the Covid-19 Pandemic," *Jurnal Obsesi : Jurnal Pendidikan Anak Usia Dini*, vol. 7, no. 4, pp. 4386–4396, Aug. 2023, doi: 10.31004/obsesi.v7i4.3722.

[28]   D. G. H. Divayana, I. P. W. Ariawan, and M. K. W. Giri, "Cipp-saw application as an evaluation tool of e-learning effectiveness," *International Journal of Modern Education and Computer Science*, vol. 13, no. 6, pp. 42–59, Dec. 2021, doi: 10.5815/ijmecs.2021.06.05.

[29]   T. D. T. Pham *et al.*, "Applying Microsoft Forms Software in Formative Assessment in Teaching Mathematical Concepts for 7th Graders in Vietnam," *Vietnam Journal of Education*, pp. 203–217, Dec. 2024, doi: 10.52296/vje.2024.400.

[30]   A. Rasul, R. Sond, and H. Sutirta, "The Microsoft form as a scaffolding-based geometry learning media in online learning," *Journal for the Mathematics Education and Teaching Practices*, vol. 3, no. 1, pp. 1–12, 2022.