VIDYA KARYA

Is The CIPP Model More Effective in Evaluating Counseling Services in The Merdeka Curriculum?

Muhammad Andri Setiawan*, Nina Permata Sari, Eklys Cheseda Makaria, Novitawati, Salwaa Alma Rizani, and Muhammad Refli

Guidance and Counseling Study Program, Faculty of Teacher Training and Education
Universitas Lambung Mangkurat, Banajrmasin, Indonesia
*andri.bk@ulm.ac.id

Abstract. This study aimed to evaluate the effectiveness of the CIPP (Context, Input, Process, Product) evaluation model in assessing Guidance and Counseling (GC) services within the Merdeka Curriculum at Vocational High School 3 and Vocational High School 4 in Banjarmasin. A quantitative method with a comparative approach was employed. Data were collected using questionnaires completed by students from both schools, and the data analysis involved a t-test to examine differences between the schools. The findings revealed that Vocational High School 4 achieved higher evaluation scores across all CIPP components compared to Vocational High School 3 Banjarmasin. Significant differences were observed in the context, input, process, and product components, with p-values <0.05. These results indicated that GC services at Vocational High School 4 Banjarmasin were more effective in implementing the Merdeka Curriculum. The study concluded that the CIPP model is an effective evaluation tool for identifying the strengths and weaknesses of GC service programs. The implications of this research suggest that improvements in the input and process components could enhance the quality of GC services at Vocational High School 3 Banjarmasin.

Keywords: CIPP model; counseling; evaluation; merdeka curriculum; vocational high school

How to cite: Setiawan, M. A., Sari, N. P., Makaria, E. C., Novitawati, N., Rizani, S. A., & Refli, M. (2024). Is the cipp model more effective in evaluating counseling services in the merdeka curriculum?. *Vidya Karya*, 39(2), 152-160.

INTRODUCTION

Implementing the Merdeka Curriculum in Indonesia aims to provide greater flexibility and autonomy for students and teachers in teaching and learning (Fauzan et al., 2023; Harto et al., 2023; Naely et al., 2022). One crucial aspect of this curriculum is the Guidance and Counseling (GC) services, designed to assist students in addressing personal, social, academic, and career challenges. According to Nehe et al. (2024), effective GC services can significantly contribute to the holistic development of students' competencies. Evaluating the effectiveness of these services is essential, and various evaluation models have been employed to assess their quality and impact. One commonly utilized model is the CIPP (Context, Input, Process, Product) model, which provides a comprehensive framework for evaluating multiple aspects of a program (Susanto et al., 2024; Toosi et al., 2021).

Implementing the Merdeka Curriculum has demonstrated increased student participation in various extracurricular activities (Andi, 2024; Haq, 2024; Prahastina et al., 2024). Data indicate that 75% of students engaged in at least one extracurricular activity, representing a significant improvement compared to the previous year (Khairi et al., 2023). This finding suggests that a more flexible curriculum allows students to explore their interests beyond the classroom (Yuliyanto & Rahmanto, 2023).

In the context of Guidance and Counseling (GC) services, adequate support is essential to helping students adapt to the changes they face. Effective GC services not only assist students in overcoming academic challenges but also contribute to character development and the enhancement of social skills (Supiah, 2023). Research by Budiman et al. (2022) revealed that students who receive support from GC services tend to exhibit higher learning motivation and are better equipped to manage academic pressure (Syamsiar et al., 2023).

Therefore, it is necessary to evaluate school-based GC programs using the CIPP model. Research by Djamaluddin et al. (2024) indicated that applying the CIPP model in evaluating GC programs at vocational high schools provides deeper insights into the effectiveness of the services offered. The findings of this study demonstrate that the CIPP model helps identify factors influencing the success of GC services (Della, 2024) and offers recommendations for improvement (Abdal et al., 2022).

Although the CIPP model is widely recognized as an effective evaluation tool, gaps remain in research regarding its application in evaluating GC services in the era of the Merdeka Curriculum (Fadila, 2024). Furthermore, no studies have yet compared the effectiveness of GC service evaluations using the CIPP model across schools with diverse backgrounds. Research by Sari and Setiawan (2023) highlights that a comprehensive understanding of context, input, process, and product in GC services can enhance program effectiveness. Therefore, there is a pressing need for more in-depth studies on the effectiveness of the CIPP model in evaluating GC services within vocational education settings (Abshor, 2021).

This study aims to answer the central question: Is the CIPP model more effective in evaluating GC services within the Merdeka Curriculum? The research was conducted at Vocational High School 3 Banjarmasin and Vocational High School 4 Banjarmasin, as both schools implement diverse GC services within the context of the Merdeka Curriculum. This study is significant in evaluating the effectiveness of GC services at both schools, particularly in identifying differences in the implementation of GC services based on the CIPP Evaluation Model used in this curriculum.

The GC services at Vocational High School 3 and Vocational High School 4 Banjarmasin show differing results. Based on preliminary observations, Vocational High School 4 has stronger support in terms of policies, teacher training, and better facilities to implement the GC program. At Vocational High School 3, although GC services are implemented, limitations in resources and support pose challenges that affect the program's effectiveness.

By analyzing evaluation data from both schools, this study aims to identify differences in the implementation of GC services and assess the CIPP model's effectiveness in measuring the GC program's success in vocational education settings (Rosalina, 2024). This research is expected to significantly contribute to developing more effective GC services that are relevant to student's needs in the era of the Merdeka Curriculum.

METHOD

This study was conducted at two vocational high schools in Banjarmasin, namely Vocational High School 3 Banjarmasin and Vocational High School 4 Banjarmasin, which had different characteristics and implementations of GC services.

This study used a comparative quantitative approach to compare the evaluation results of GC services using the CIPP model at Vocational High School 3 Banjarmasin and Vocational High School 4 Banjarmasin. This approach allowed the researcher to assess differences in the effectiveness of GC programs between the two schools and to measure the extent to which the CIPP model could be used to evaluate GC services within the context of the Merdeka Curriculum.

This study was conducted over six months, from January to June 2024. The respondents were students from Vocational High School 3 Banjarmasin and Vocational High School 4 Banjarmasin, who had received GC services while implementing the Merdeka Curriculum. The population in this study consisted of all eleventh-grade students at both schools, with a sample selected using random sampling. The total sample size was 600 students, with 300 students from Vocational High School 3 Banjarmasin and 300 from Vocational High School 4 Banjarmasin.

Data were collected through the following methods: (1) a questionnaire, which served as the primary instrument in this study. The questionnaire was designed based on the components of the CIPP model (Context, Input, Process, Product). It contained statements measured using a Likert scale, assessing students' perceptions of GC services at their schools. Each statement in the questionnaire was scored on a scale from 1 to 5, where 1 meant "strongly disagree" and 5 meant "strongly agree." These scores reflected how students assessed the effectiveness of GC services based on context, resources, implementation process, and the outcomes produced by the program; (2) semi-structured interviews were conducted with GC teachers at each school to gather additional information regarding the implementation of the GC program and how these services were adapted to the Merdeka Curriculum; and (3) documentation in the form of secondary data collection was carried out through the evaluation reports of the GC programs at Vocational High School 3 Banjarmasin and Vocational High School 4 Banjarmasin to ensure the data collected was valid and could be comprehensively compared.

The collected data were analyzed using the following techniques: (1) descriptive analysis was conducted to provide an overview of the evaluation of GC services at both schools based on the CIPP components (context, input, process, and product). Descriptive analysis was performed on the mean, standard deviation, and data distribution from the questionnaire; (2) an independent t-test was conducted to examine significant differences between the two schools. The independent t-test was applied to each CIPP component. This test was used to test the hypothesis that there were significant differences in the evaluation results of GC services between Vocational High School 3 Banjarmasin and Vocational High School 4 Banjarmasin, and (3) Levene's test was performed to ensure the equality of variances between the two groups (Vocational High School 3 Banjarmasin and Vocational High School 4 Banjarmasin) before performing the t-test. If Levene's test indicated heterogeneous variances, adjustments were made to the statistical test used.

RESULTS AND DISCUSSION

The analysis was conducted based on the four main components of the CIPP model:

Context, Input, Process, and Product. The evaluation results for each component were compared between the two schools and interpreted to address the research question regarding the effectiveness of the CIPP evaluation model.

Evaluation Results Based on Context
Based on the descriptive analysis presented in Table 1, the mean score for the context evaluation at Vocational High School 4 Banjarmasin (3.5587 ± 0.37774) was higher than that at Vocational High School 3 Banjarmasin (2.7310 ± 0.58788). This indicated that Vocational High School 4 Banjarmasin achieved better context evaluation results compared to Vocational High School 3 Banjarmasin.

Table 1 Descriptive analysis of evaluation based on context

School	N	Mean	Std. Deviation
Vocational High Scool 3 Banjarmasin	300	2.7310	0.58788
Vocational High Scool 4 Banjarmasin	300	3.5587	0.37774

The t-test results in Table 2 also showed a significant difference between the two schools, with a significance value of 0.000 (<0.05). Furthermore, the variances between the two schools were not homogeneous, as indicated by Levene's test results (0.000).

Table 2 T-test analysis of evaluation based on context

	Sig f test (Levene's test)	Sig t-test
Context	0.000	0.000

These findings indicated that Vocational High School 4 Banjarmasin had stronger contextual support for implementing GC services than Vocational High School 3 Banjarmasin. This support was evident in more favourable school policies, adequate facilities, and greater parental involvement (Della, 2024). According to Kurniawan (2024), robust contextual support significantly influences the success of educational programs. At Vocational High School 4, intensive training programs for GC teachers and collaborations with external institutions supported the development of GC services. Conversely, at Vocational High School 3, limited resources and insufficient policy support posed challenges to effectively implementing GC services (Basuki & Anriani, 2023). These findings highlighted that strong contextual factors could be pivotal in determining the success of GC program implementation (Dalmia & Alam, 2021).

Evaluation Results Based on Input
For the input component, the descriptive analysis in Table 3 showed that Vocational High School 4 Banjarmasin also had a higher mean score (2.9539 ± 0.41473) compared to Vocational High School 3 Banjarmasin (2.7163 ± 0.57640).

Table 3 Descriptive analysis of evaluation based on input

School	N	Mean	Std. Deviation
Vocational High School 3 Banjarmasin	300	2.7163	0.57640
Vocational High School 4 Banjarmasin	300	2.9539	0.41473

The t-test showed a significant difference with a significance value of 0.000, and Levene's test also indicated non-homogeneous variances (0.000), as shown in Table 4.

Table 4 T-test analysis of evaluation based on input

	Sig f test (Levene's test)	Sig t-test
Input	0.000	0.000

The significant difference in the input component between the two schools indicated that Vocational High School 4

Banjarmasin had better resources, such as the availability of trained GC teachers, supporting facilities, and more comprehensive GC materials. This is important because good input factors will affect the process and outcomes of GC services (Nehe, 2024). These findings reinforce the view that the input component is one of the key factors in successfully implementing GC services (Khairi et al., 2023; Syamsiar et al., 2023).

Evaluation Results Based on Process
In the process aspect, Table 5 showed that the mean evaluation score for the process at Vocational High School 4 Banjarmasin (3.4289 ± 0.31044) was slightly higher than at Vocational High School 3 Banjarmasin (3.3027 ± 0.35756). Although this mean difference was not as large as in the previous components, the t-test results showed a significant difference with a significance value of 0.000.

Table 5 Descriptive analysis of evaluation based on process

School	N	Mean	Std. Deviation
Vocational High School 3 Banjarmasin	300	3.3027	0.35756
Vocational High School 4 Banjarmasin	300	3.4289	0.31044

Levene's test also indicated non-homogeneous variances (0.003), as shown in Table 6.

Table 6 T-test analysis of evaluation based on process

	Sig f test (Levene's test)	Sig t-test
Input	0.003	0.000

The process evaluation showed that although both schools implemented GC services fairly well, Vocational High School 4 Banjarmasin demonstrated an advantage in executing GC procedures and activities. This could have been caused by differences in GC teacher training or student involvement in the counselling process (Farid, 2024). An effective process will support the achievement of the goals of GC services, as explained in the CIPP theory, which emphasizes the importance of evaluating how services are delivered (Sari et al., 2023).

Evaluation Results Based on Product
In the product component, the evaluation results showed that Vocational High School 4 Banjarmasin had a higher mean score (3.3908 ± 0.68881) compared to Vocational High School 3 Banjarmasin (2.8017 ± 0.83771), as shown in Table 7.

Table 7 Descriptive analysis of evaluation based on product

School	N	Mean	Std. Deviation
Vocational High School 3 Banjarmasin	300	2.8017	0.83771
Vocational High School 4 Banjarmasin	300	3.3908	0.68881

The t-test results indicated a significant difference with a significance value of 0.000 and non-homogeneous variances (Levene's test: 0.000), as shown in Table 8.

Table 8 T-test analysis of evaluation based on product

	Sig f test (Levene's test)	Sig t-test
Product	0.000	0.000

The counselling services' products at Vocational High School 4 Banjarmasin were better than those at Vocational High School 3 Banjarmasin, reflecting more positive outcomes such as improved student achievement, mental well-being, and problem-solving skills. The research by Nehe et al. (2024) and Khairi et al. (2023) emphasized that the use of the CIPP model in evaluating counselling services helps identify the components contributing to service success, as seen in

the improvement of students' well-being and their ability to face academic and personal challenges (Farid, 2024).

Furthermore, Nehe et al. (2024) highlighted the importance of product-based evaluation to ensure the effectiveness of services that have a direct impact on students. These studies emphasize that product-based evaluation provides an overview of a program's long-term impact (Subhaktiyasa et al., 2022), as observed at Vocational High School 4 Banjarmasin.

Overall Evaluation Results
Overall, Table 9 shows that the overall evaluation results also indicate the superiority of Vocational High School 4 Banjarmasin, with a mean score of 3.3337 ± 0.25420, compared to Vocational High School 3 Banjarmasin, which has a mean score of 2.8887 ± 0.49362.

Table 9 Descriptive analysis of overall evaluation results

School	N	Mean	Std. Deviation
Vocational High School 3 Banjarmasin	300	2.8887	0.49362
Vocational High School 4 Banjarmasin	300	3.3337	0.25420

The t-test results again show a significant difference with a significance value of 0.000, and Levene's test shows non-homogeneous variance (0.000), as shown in Table 10.

Table 10 T-test for overall evaluation results

	Sig f test (Levene's test)	Sig t-test
Input	0.000	0.000

Based on the results of Levene's test presented in Table 2, Table 4, Table 6, and Table 8, it was found that the variances between the two schools (Vocational High School 3 Banjarmasin and Vocational High School 4 Banjarmasin) were not homogeneous, as

indicated by a significance value of <0.05. This lack of homogeneity is important in statistical testing, particularly when comparing the two groups using difference tests such as the independent t-test.

On the other hand, when the variances were not homogeneous, the standard t-test was not entirely appropriate. To address this issue of non-homogeneity, a more suitable approach was to use Welch's t-test, which was a modification of the independent t-test and was designed to handle unequal variances between groups. Welch's t-test allowed for testing the differences between the two groups even when their variances were not homogeneous, ensuring more accurate results that met the statistical testing assumptions.

Therefore, although it was intuitive to assume that two different schools, such as Vocational High School 3 and Vocational High School 4 Banjarmasin, would have differences, given the varying conditions and contexts, the purpose of using statistical tests was not simply to show that there were differences, but to measure whether these differences were statistically significant. This testing provided a more scientific and objective justification than relying solely on general observations.

Discussion of Findings
The evaluation results presented in this study showed that Vocational High School 4 Banjarmasin overall had better outcomes across all CIPP components than Vocational High School 3 Banjarmasin. Although both schools were not homogeneous in their variances, the use of appropriate statistical tests still provided a stronger and more scientific understanding of the differences in counselling services' effectiveness.

The results showed that Vocational High School 4 Banjarmasin consistently excelled in all evaluation components—context, input, process, and product. This

outcome aligned with the theory that strong contextual support, such as supportive policies, adequate facilities, and parental involvement, forms a crucial foundation for the success of educational programs (Kurniawan, 2024; Dalmia & Alam, 2021). These findings also supported the view that quality input, such as trained educators and adequate facilities, directly impacts the effectiveness of the service process and the outcomes achieved by students (Khairi et al., 2023).

Vocational High School 4 Banjarmasin's excellence in all evaluation components (context, input, process, and product) provided empirical evidence that the CIPP model could comprehensively identify a program's strengths and weaknesses.

In addition, the significant differences between the two schools reinforced the view that supportive policies, availability of quality resources, and good processes strongly influenced the outcomes of counselling services. While the non-homogeneity of variances had to be noted, the use of appropriate statistical tests, such as Welch's t-test, ensured that the results remained valid and reliable. Therefore, this study affirmed that the CIPP evaluation model was a reliable tool for assessing the effectiveness of counselling programs.

CONCLUSION

This study had several limitations that needed to be considered when interpreting the results. First, the heterogeneity of variance between the two sample groups indicated the potential for bias due to differences in external contexts, such as students' socioeconomic characteristics, regional policy support, or educational budget allocation. This could have affected the evaluation results and provided disproportionate advantages to certain schools.

Second, this study only focused on two schools, Vocational High School 3 and Vocational High School 4 Banjarmasin, which may not have been representative enough to reflect the conditions of GC services at the national or regional level. Generalizing the research results to other schools with different contexts should have been done cautiously.

Third, the analysis relied solely on quantitative data from the CIPP model without delving deeper into qualitative aspects, such as interviews with GC teachers, students, or parents. A qualitative approach could have provided deeper insights into how certain factors contributed to the effectiveness of GC services.

Despite these limitations, this study contributed to developing research on the CIPP evaluation model in the context of GC services, particularly in implementing the Merdeka Curriculum. The study showed that the CIPP model could be used as a comprehensive evaluation tool to identify various aspects of educational services, including context, input, process, and product.

Specifically, this study's findings provided guidance for schools in utilizing evaluation results to improve the quality of GC services. For example, context factors such as supportive school policies, input factors like adequate GC teacher training, and process factors such as student involvement in counselling were highlighted as important elements that could serve as a basis for decision-making.

On a broader scale, this study reinforced the literature on the use of the CIPP model in Indonesia, particularly in vocational education settings in the era of the Merdeka Curriculum. The model proved to be relevant in helping schools align their GC programs with students' diverse needs, whether in academic development, career guidance, or mental well-being.

REFERENCES

Abdal, N. M., Suwahyu, I., & Sulaiman, D. R. A. (2022). Pengembangan instrumen evaluasi program asistensi mengajar di satuan pendidikan mandiri program studi pendidikan teknik informatika dan komputer unm untuk mitra sekolah dengan model cipp. *Information Technology Education Journal*, *1*(2), 42-45.

Abshor, M. (2021). Evaluasi program visiting teacher pendidikan agama islam di wilayah perbatasan. Tesis. Universitas Islam Negeri Syarif Hidayatullah

Andi, W. (2024). *Implementasi manajemen kurikulum merdeka di smk negeri 7 bandar lampung*. Skripsi. Universitas Islam Negeri Raden Intan Lampung.

Basuki, B., & Anriani, N. (2023). Evaluation of accounting learning models: implementation of context, input, process, and product (cipp) evaluation on problem-based learning models. *Edunesia: Jurnal Ilmiah Pendidikan*, *4*(2), 711-721.

Budiman, C., Badrujaman, A., & Wahyuni, E. (2022). Evaluasi program bimbingan dan konseling bidang sosial dengan teknik Context, Input, Proses, Produk (CIPP) di sekolah menengah kejuruan. *Jurnal Konseling dan Pendidikan*, *10*(2), 354-363.

Dalmia, D., & Alam, F. A. (2021). Evaluasi program model context dan input dalam bimbingan konseling. *Jurnal Bimbingan Konseling dan Psikologi*, *1*(2), 111-124.

Della, P. K. (2024). *Evaluasi program pendidikan karakter di mts guppi natar lampung selatan*. (Doctoral dissertation, UIN RADEN INTAN LAMPUNG).

Djamaluddin, S., Rahabav, P., Sahalessy, A., & Rumfot, S. (2024). Manajemen Pendidikan Karakter di SMA Negeri 13 Ambon. *EDUKASIA: Jurnal Pendidikan dan Pembelajaran*, *5*(1), 1689-1700.

Fadila, M. D. (2024). Evaluasi program bimbimbingan konseling menggunakan model countenance stake di smp al fath cirendeu. Universitas Islam Negeri Syarif Hidayatullah

Farid, A. (2024). *Manajemen internalisasi nilai pendidikan karakter berbasis multikultural di smp labschool jakarta*. (Doctoral Dissertation, Universitas Negeri Jakarta).

Fauzan, F., Ansori, R. A. M., Dannur, M., Pratama, A., & Hairit, A. (2023). The implementation of the merdeka curriculum (independent curriculum) in strengthening students' character in Indonesia. *Aqlamuna: Journal of Educational Studies*, *1*(1), 136-155.

Harto, M., Misbah, M., Purwasih, D., Umar, F., & Muhammad, N. (2023). Bibliometric analysis of research trends on differentiated learning in merdeka curriculum. *Vidya Karya*, *38*(1), 14-24.

Haq, H. (2024). Evaluation of the implementation of the merdeka belajar curriculum in secondary schools in the digital era. *International Journal of Post Axial: Futuristic Teaching and Learning*, 215-228.

Khairi, A., Jalinus, N., Ambyar, A., & Waskito, W. (2023). Evaluation of the implementation of independent learning-independent campus (mbkm) curriculum by using the cipp evaluation model. *Scaffolding: Jurnal Pendidikan Islam dan Multikulturalisme*, *5*(2), 543-559.

Kurniawan, M. D. (2024). *Hubungan antara supervisi kepala sekolah terhadap motivasi berprestasi dengan kinerja guru madrasah ibtidayah kecamatan wanaraya kabupaten barito kuala* (Doctoral dissertation, Universitas Islam Kalimantan MAB).

Naely, U., Asha, L., Azwar, B., Warlizasusi, J., & Sumarto, S. (2022). Evaluasi pelaksanaan asesmen nasional tahun 2021. *Journal Of Administration and Educational Management (ALIGNMENT)*, *5*(2), 299-306.

Nehe, A. (2024). *Evaluasi Program Bimbingan Dan Konseling Komprehensif Dengan Menggunakan Model Cipp Di Smk Swasta Pembda Nias*. Doctoral dissertation. Universitas Kristen Satya Wacana.

Nehe, A., Satyawati, S. T., & Dwikurnaningsih, Y. (2024). Evaluation of comprehensive guidance and counseling program using the CIPP model. *Inovasi Kurikulum*, *21*(3), 1633-1648.

Prahastina, L., Indriayu, M., & Matsuri, M. (2024). Exploring the effectiveness of the Merdeka curriculum in promoting effective learning practices. *JPPI (Jurnal Penelitian Pendidikan Indonesia)*, *10*(3), 111-121.

Rosalina, T. (2024). Penguatan kompetensi siswa melalui program magang pada dunia usaha dan dunia industri. *Al-Muttaqin: Jurnal Studi, Sosial, dan Ekonomi*, *5*(2), 173-184.

Sari, N. P., & Setiawan, M. A. (2023). Evaluation analysis based on the cipp model vocational high school guidance and counseling program: Expert perspective, guidance and counseling teacher, and subject teacher. *Eurasian Journal of Educational Research*, *106*(106), 29-49.

Sari, N. P., Setiawan, M. A., Makaria, E. C., Dinanty, N. S., & ... (2023). *Pengembangan model evaluasi cipp untuk program bimbingan dan konseling pada kurikulum merdeka tingkat smk di wilayah lahan basah*.

Subhaktiyasa, P. G., Ratnaya, I. G., Nazim, M., Bayangkari, B., & Suprapta, M. A. (2022). The evaluation of career centre management at health colleges: a cipp model approach. *Jurnal Basicedu*, *6*(5), 8069-8077.

Supiah, M. P. (2023). *Ilmu Manajemen Pendidikan Islam*. Selat Media.

Susanto, A., Kardena, A., & Elpayani, E. (2024). Evaluation of the cipp model on the merdeka curriculum at smp 2 palembayan. *EDUMALSYS Journal of Research in Education Management,* 2(2).

Syamsiar, H., Muzakki, M., Ratnaya, I. G., & Widiana, I. W. (2023). Evaluasi Kurikulum Merdeka Belajar Berbasis CIPP. *Journal of Education and Instruction (JOEAI)*, *6*(2), 536-544.

Toosi, M., Modarres, M., Amini, M., & Geranmayeh, M. (2021). Context, Input, Process, and Product Evaluation Model in medical education: A systematic review. *Journal of education and health promotion*, *10*(1), 199.

Yuliyanto, R., & Rahmanto, A. A. (2023). Literature review evaluation of fieldwork practice programs in vocational high schools in indonesia. *Jurnal Hurriah: Jurnal Evaluasi Pendidikan dan Penelitian*, *4*(4), 433-443.