Copyright © 2026 The Author(s)
Didaktika: Jurnal Kependidikan Vol. 15 No. 3, Agustus 2026

Evaluating a BLUD-Based Teaching Factory in Motorcycle Engineering: A CIPP Study at an Indonesian Vocational High School

A. Jufrizal[1], Zainal Arifin[2]
[1,2]Universitas Negeri Yogyakarta, Indonesia

[1]ajufrizal.2024@student.uny.ac.id
DOI: https://doi.org/10.58230/27454312.5141

Abstract

Teaching Factory programs seek to align vocational learning with authentic industrial work, while Indonesia's Regional Public Service Agency (Badan Layanan Umum Daerah, BLUD) arrangement provides public vocational schools with greater operational and financial flexibility. This study evaluated a BLUD-based Teaching Factory in the Motorcycle Engineering program at SMK Negeri 2 Tanjungpandan using the Context, Input, Process, Product (CIPP) model. The evaluation used descriptive quantitative data collected from December 2025 to March 2026. Role-specific questionnaires were completed by 62 students, seven productive-subject teachers, seven BLUD managers, and five industry representatives (N = 81). Two observers also assessed implementation, and a 40-item document checklist was reviewed. The overall questionnaire mean was 4.23 on a five-point scale (84.60%, very good). Context scored 4.21 (84.20%), Input 4.19 (83.80%), Process 4.24 (84.80%), and Product 4.28 (85.60%). The strongest indicators were occupational safety and standard operating procedure implementation (M = 4.55), industry partnership support (M = 4.39), and perceived impact on learning and students (M = 4.36). The lowest indicators concerned administrative and financial system readiness (M = 4.00), learning management and evaluation (M = 4.04), and service or production quality (M = 4.04). Observation and documentation results were 81.91% and 85.00%, respectively. The program therefore showed strong relevance, implementation, and perceived outcomes, but the cross-sectional, single-site design does not establish causal improvement in work readiness. Priority actions are to modernize facilities, strengthen BLUD administration and monitoring, and add longitudinal graduate and employer outcome measures.

Keywords: *BLUD, CIPP evaluation, Teaching Factory, vocational education, work readiness.*

Introduction

Vocational education is expected to prepare learners for occupational performance in rapidly changing workplaces. In Indonesia, this expectation is especially salient because vocational high schools must reconcile national curriculum requirements, local labor-market conditions, and industrial standards. Persistent criticism of school-to-work mismatch has encouraged models that place authentic work, industrial quality requirements, and customer service inside the learning process. Teaching Factory is one of the most prominent responses because it organizes learning around real or simulated production and service activities rather than isolated classroom exercises (Suharno et al., 2020; Diwangkoro & Soenarto, 2020).

A Teaching Factory is more than a workshop equipped with machines. It requires a coherent relationship among curriculum, work orders, instructors, students, industrial

partners, quality standards, and assessment. When these elements are aligned, students can connect theoretical knowledge with occupational routines, develop technical competence, and learn communication, teamwork, discipline, and responsibility. Studies in Indonesian vocational settings have reported positive associations between Teaching Factory participation and occupational or work readiness (Khoiron, 2016; Dewi & Sudira, 2018; Prianto et al., 2021). Recent studies likewise emphasize the value of authentic service-production units and industry-responsive learning for graduate readiness (Rosidah & Sutirman, 2023; Imran et al., 2024).

However, the presence of a Teaching Factory does not guarantee that the program functions as an integrated learning system. Implementation studies identify recurring problems in staff competence, equipment currency, production management, monitoring, curriculum alignment, and the continuity of industrial partnerships (Kautsar et al., 2022; Chasovy et al., 2024). These constraints are interdependent. For example, a school may possess adequate equipment but lack a reliable order-management system, or it may have an industrial partner without translating current workplace standards into modules, job sheets, and assessments. A comprehensive evaluation must therefore examine not only outcomes but also the conditions and processes that produce them.

The Context, Input, Process, Product (CIPP) model is appropriate for this purpose because it links program needs and goals with resources, implementation, and results. Context evaluation asks whether the program responds to relevant needs and policies; Input evaluation examines strategies and resources; Process evaluation assesses whether planned activities are implemented as intended; and Product evaluation considers intended and unintended outcomes. The model is decision-oriented: its value lies not only in judging a program but also in informing continuation and improvement (Stufflebeam & Shinkfield, 2007; Frye & Hemmer, 2012). Reviews of evaluation models and vocational applications further support CIPP as a systematic framework for complex educational programs (Nouraey et al., 2020; Ratnaya et al., 2022).

In the Indonesian public-school context, the governance arrangement is also consequential. A Regional Public Service Agency, known as Badan Layanan Umum Daerah (BLUD), allows designated public institutions to manage services and finances with greater flexibility while maintaining public accountability under Minister of Home Affairs Regulation No. 79 of 2018. For vocational schools, this flexibility can help Teaching Factory units purchase consumables, maintain equipment, manage service income, and respond to customer orders more rapidly. At the same time, BLUD status introduces demands for planning, documentation, financial controls, transparency, and professional service management (Ministry of Home Affairs of the Republic of Indonesia, 2018).

Research on BLUD vocational schools indicates that financial flexibility can strengthen industry-based learning when governance functions, service standards, and accountability mechanisms are coherent (Hariyanto, 2021; Wibawa et al., 2023). Yet Teaching Factory research frequently focuses on learning design, competence, or work readiness, whereas BLUD research often focuses on institutional or financial management. The intersection between pedagogical implementation and BLUD governance remains less visible, particularly in school-level evaluations that combine perspectives from students, teachers, managers, and industry.

Motorcycle Engineering provides a demanding setting for such an evaluation. Students must diagnose faults, use tools safely, follow service procedures, communicate with customers, manage time and parts, document work, and meet quality expectations. These tasks require a learning environment that connects technical practice with the organizational

Copyright © 2026 The Author(s)

routines of an actual service unit. A BLUD-based Teaching Factory could support this integration, but its effectiveness should be assessed through evidence about relevance, resources, operations, and perceived outcomes rather than assumed from its formal designation.

This study therefore evaluated the BLUD-based Teaching Factory in the Motorcycle Engineering program at SMK Negeri 2 Tanjungpandan. The evaluation addressed four questions: (1) How relevant are the program goals, policies, and partnerships to school and industry needs? (2) How adequate are the human, material, curricular, partnership, administrative, and financial inputs? (3) How well are learning, production or service operations, safety procedures, BLUD management, and monitoring implemented? and (4) What outcomes are perceived for students, learning, work readiness, and the school? The study's contribution is an integrated CIPP profile that combines multi-stakeholder questionnaires with observation and documentation to identify both program strengths and concrete improvement priorities in a BLUD-governed vocational service setting.

Method

This was a program-evaluation study using a descriptive quantitative approach and the CIPP model (Creswell & Creswell, 2018). The evaluation was conducted at SMK Negeri 2 Tanjungpandan, Belitung Regency, Bangka Belitung Islands Province, Indonesia. The focal program was a Motorcycle Engineering Teaching Factory operated within the school's BLUD governance arrangement. Field activities were conducted from December 2025 to March 2026 and covered preparation, data collection, analysis, and reporting.

The evaluation was designed to describe the condition of the program at the time of data collection. It did not use a pretest-posttest comparison, a control group, random assignment, or longitudinal follow-up. Accordingly, the product component represents stakeholder judgments of program outcomes and perceived work readiness; it should not be interpreted as a causal estimate of improvement.

The questionnaire dataset comprised 81 respondents: 62 Year 11 and Year 12 Motorcycle Engineering students, seven productive-subject teachers, seven BLUD managers, and five industry representatives. The two sampling terms reported in the thesis describe sequential and complementary decisions rather than competing procedures. First, purposive eligibility defined the student sampling frame as Year 11 or Year 12 learners who had completed the Teaching Factory program in full for at least one year. Second, because the resulting accessible frame contained 62 eligible students, every eligible student was included; this was saturated sampling, or a census, within the purposively defined frame. Thus, the student findings describe all accessible program-experienced students, not the entire school population. The non-student respondents were role-based key informants: productive teachers and BLUD managers directly responsible for the program, and industry representatives with direct partnership experience. Two observers provided structured implementation ratings. Documentary evidence included curriculum and learning documents, standard operating procedures, BLUD administration, Teaching Factory reports, service records, and field photographs.

Table 1. Questionnaire participants and analytical contribution

Evidence source	n	Instrument or record	Primary contribution
Students	62	Likert questionnaire	Learning experience, implementation, skills, and readiness
Productive teachers	7	Likert questionnaire	Goals, instructional inputs, processes, and outcomes
BLUD managers	7	Likert questionnaire	Governance, finance, operations, and sustainability
Industry representatives	5	Likert questionnaire	Relevance, partnerships, standards, and readiness

Role-specific questionnaires used a five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). Indicators were organized around the four CIPP components. Context covered industry alignment, clarity of program goals, partnership policy, and BLUD-policy relevance. Input covered teachers and human resources, facilities, modules and job sheets, industry support, administrative and financial systems, and curriculum or competency readiness. Process covered Teaching Factory instruction, learning management and assessment, BLUD activities and finance, monitoring, service quality, standard operating procedures and occupational health and safety (OHS), production or service management, teamwork, and discipline. Product covered hard skills, soft skills, work readiness, learning impact, school impact, and program sustainability.

Content validity was assessed before field administration by three experts with university, vocational-education, and school-supervisory expertise. They reviewed alignment between the CIPP indicators and questionnaire items, redundancy, wording clarity, indicator coverage, and the adequacy of observation subindicators. Their written recommendations included removing duplicated items, aligning the number of indicators with the number of questions, refining student items concerning facilities, adding observation detail concerning teacher certification, methodology, and experience, and revising ambiguous wording. The instruments were revised in response to this feedback. Construct-related evidence was then examined in a pilot with 30 respondents outside the study sample using Pearson item-total correlations at the 5% significance level. Six of the 45 pilot items fell below the critical correlation and were flagged for revision or removal. The thesis reports a Cronbach's alpha of .926 for the pilot questionnaire set, exceeding the .70 criterion. However, the thesis does not provide a quantitative content-validity coefficient, post-revision item coefficients, or reliability estimates for each role-specific form. The available evidence therefore supports a documented expert-review process and high overall pilot internal consistency, but it does not demonstrate that every CIPP subscale was equally valid and reliable.

Questionnaire data were summarized using means and percentages. Percentage achievement was calculated as the component mean divided by the maximum score of five and multiplied by 100. The thesis defined a mean of at least 3.50 as the program-success threshold and converted component scores into qualitative categories. The qualitative labels reported in this article reproduce the source thesis' conversion scheme. Indicator-level means were examined to identify the strongest practices and the most actionable constraints.

Copyright © 2026 The Author(s)

Observation achievement was calculated from the two observers' combined score relative to the maximum possible checklist score, while documentation achievement represented the proportion of required records available. Cross-source triangulation was conducted through structured comparison rather than by pooling unlike measures. The overall questionnaire percentage was compared with observation and documentation percentages, and evidence was matched at the domain level when the sources addressed the same practice. Agreement was treated as convergence, while discrepancies were retained as diagnostic findings. Documentary availability was not equated with record quality or routine use, and observation of implementation was not treated as direct evidence of a gain in work readiness.

Because the source tables did not report component-level standard deviations, confidence intervals, respondent-group distributions, or item-level post-validation coefficients, this article retains the verified descriptive statistics and does not introduce inferential tests or fabricated dispersion estimates. The qualitative categories reproduce the thesis conversion scheme, but interpretation emphasizes relative indicator gaps, cross-source convergence, and the design limitations rather than treating a category label as proof of effectiveness.

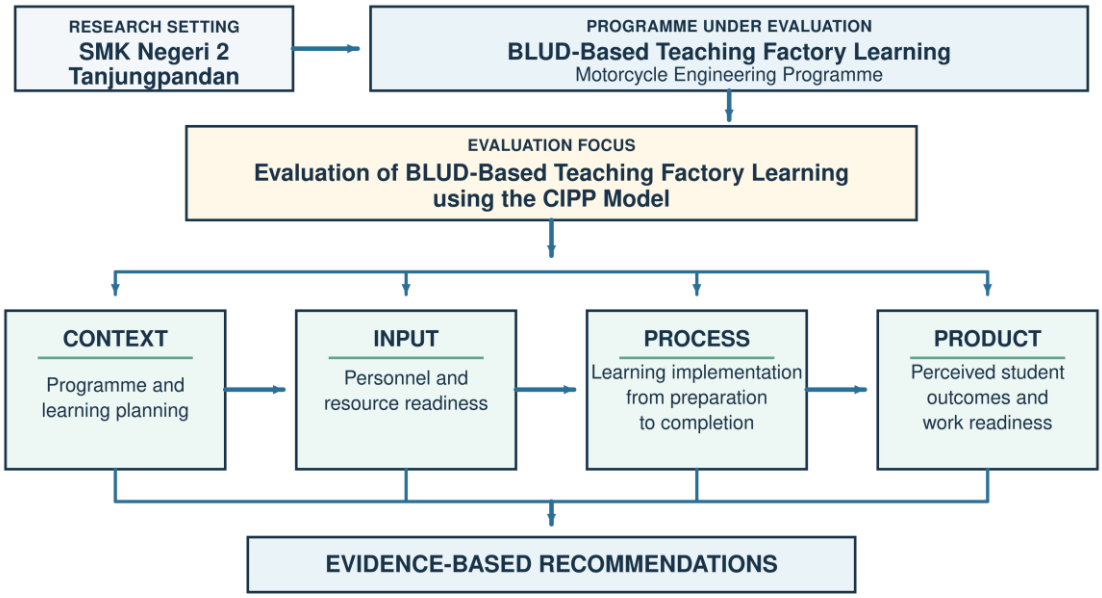


Figure 1. English adaptation of the thesis research framework for evaluating BLUD-based Teaching Factory learning using the CIPP model.

Results

Overall CIPP Profile

The BLUD-based Teaching Factory achieved an overall questionnaire mean of 4.23, equivalent to 84.60% of the maximum score and classified as very good under the thesis criteria. All four CIPP components exceeded the 3.50 success threshold. Product received the highest component mean (M = 4.28), followed by Process (M = 4.24), Context (M = 4.21), and Input (M = 4.19). The component range was only 0.09 points, showing strong upper-end concentration. This clustering is a favorable perception profile, but it is not evidence that every function operated equally well. Relative indicator gaps and cross-source comparisons are therefore more informative than the category labels alone.

Table 2. CIPP component scores for the BLUD-based Teaching Factory

Component	Mean	Score (%)	Category	Primary interpretation
Context	4.21	84.20%	Very good	Strong alignment among goals, industry needs, policy, and partnerships
Input	4.19	83.80%	Good	Adequate resources, with administrative-financial and facility priorities
Process	4.24	84.80%	Very good	Strong implementation, particularly SOP and OHS compliance
Product	4.28	85.60%	Very good	Highly rated student, learning, and school outcomes
Overall CIPP	4.23	84.60%	Very good	Program merits continuation with targeted strengthening

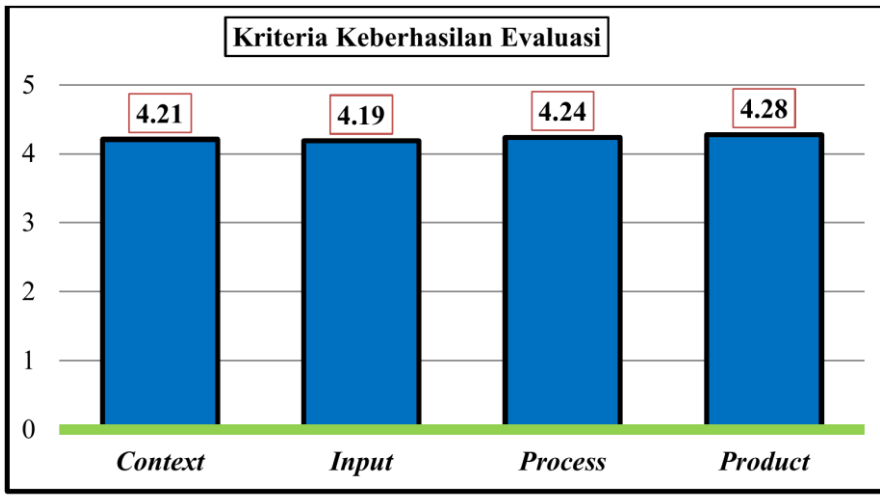


Figure 2. CIPP component mean scores reproduced from the original thesis figure; Indonesian title is retained.

Component-Level Strengths and Constraints

Within Context, the highest indicator was partnership goals and policy (M = 4.40), followed by alignment of the Teaching Factory with industry needs (M = 4.25). BLUD management policy and goals scored 4.18, and alignment of the BLUD program with school and industry needs scored 4.14. Clarity of Teaching Factory goals was the lowest Context indicator (M = 4.09), although it remained in a favorable range. These results show that external relevance and partnership structures were stronger than the consistency with which program objectives were communicated to stakeholders.

Input showed a comparable pattern. Industry support and partnership was the strongest input (M = 4.39), and curriculum and competency readiness also scored highly (M = 4.35). Teacher and human-resource readiness and module or job-sheet readiness each scored 4.16. Facilities scored 4.09. Administrative and financial system readiness was the lowest input (M =

Copyright © 2026 The Author(s)

4.00), identifying BLUD administrative capability as a priority even though the component as a whole remained successful.

Process was led by the application of standard operating procedures and occupational health and safety (M = 4.55). Production or service management and teamwork or discipline each scored 4.30; BLUD activity and financial management each scored 4.29; and program monitoring and evaluation scored 4.23. Teaching Factory instruction scored 4.20, while BLUD integration in the Teaching Factory scored 4.11. Learning management and evaluation and service or production quality were the lowest process indicators (both M = 4.04). The profile therefore combines a strong safety culture with opportunities to strengthen systematic instructional evaluation and service-quality assurance.

Product received the highest component score. Perceived impact on learning and students was the strongest indicator (M = 4.36), followed by hard skills (M = 4.35), soft skills (M = 4.32), and work readiness (M = 4.30). Impact on the school scored 4.18, while program effectiveness and sustainability scored 4.14. The results indicate strong stakeholder confidence in student and learning outcomes, but the lower sustainability score shows that maintaining resources, governance, and partnerships is a continuing rather than a completed task.

Table 3. Highest and lowest indicators within each CIPP component

Component	Highest indicator	M	Lowest indicator	M
Context	Partnership goals and policy	4.40	Clarity of Teaching Factory goals	4.09
Input	Industry support and partnership	4.39	Administrative and financial system readiness	4.00
Process	SOP and occupational health and safety	4.55	Learning management/evaluation; service quality	4.04
Product	Impact on learning and students	4.36	Program effectiveness and sustainability	4.14

Observation and Documentary Evidence

The structured observations produced a combined achievement of 81.91%, which was 2.69 percentage points below the questionnaire result of 84.60%. Documentation completeness was 85.00%, only 0.40 points above the questionnaire result. At the Process level, the observers' mean of 4.19 was close to the questionnaire Process mean of 4.24, a difference of 0.05. These comparisons show broad convergence, but not identity, across evidence sources. The strongest convergence concerned students' participation in practical work, division of tasks, application of workshop procedures, and visible industrial work routines.

Triangulation also revealed a meaningful divergence that would be hidden by the overall category. Four observed practices received a score of 3 (adequate): student involvement in BLUD administration, job-administration recording, simple financial management, and service-administration procedures. This finding qualifies the favorable Process questionnaire score and directly supports the comparatively lower ratings for administrative and financial readiness (M = 4.00) and BLUD integration in Teaching Factory learning (M = 4.11). The document

checklist confirmed that 34 of 40 expected records were available, but six were absent. Document presence therefore supports institutionalization of the program while leaving open whether each record was complete, current, and routinely used in learning and service decisions.

Field documentation illustrates an end-to-end service-learning sequence: receiving the customer and motorcycle, conducting the service work, and completing instructor-supervised technical tasks. The photographs confirm that students interacted with authentic service objects and workplace roles. They do not, however, verify diagnostic accuracy, customer satisfaction, defect or rework rates, or a change in work readiness. The visual evidence therefore corroborates the authenticity of implementation, not the quality of every service outcome or a causal student effect.

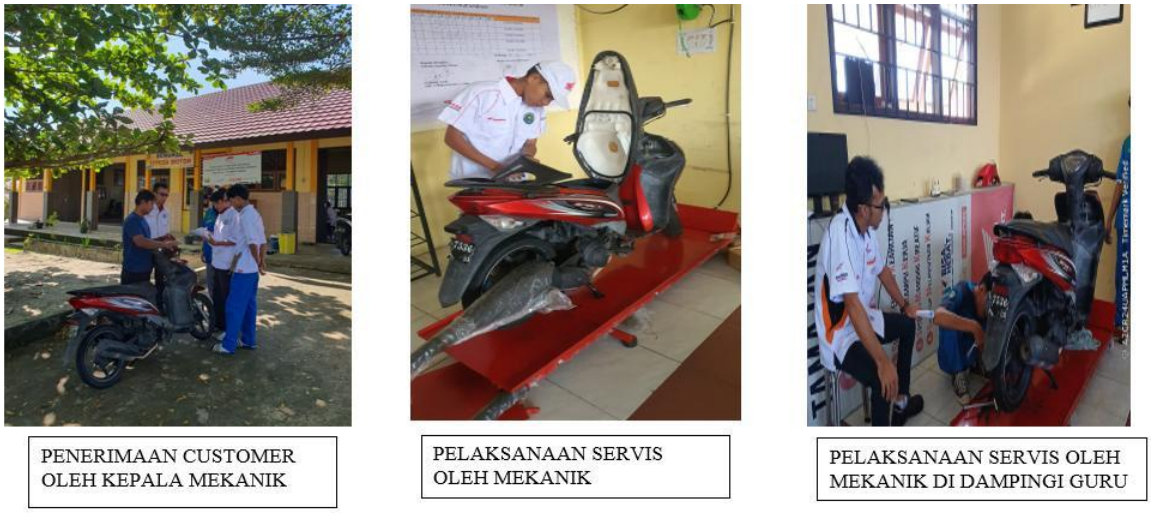


Figure 3. Teaching Factory service-learning sequence: customer reception, motorcycle servicing, and instructor-supervised servicing. Source: thesis field documentation.

Discussion

A Coherent but Uneven Program System

The evaluation presents a coherent program profile. Context, Input, Process, and Product all exceeded the success criterion, and observation and documentation broadly corroborated the favorable questionnaire ratings. The most important interpretation is not that one component was dramatically better than another, but that strong partnerships and safety-centered implementation supported highly rated student outcomes while administrative, financial, monitoring, and service-quality systems remained less mature. CIPP is useful here because these lower scores are not treated as isolated defects; they identify conditions that could eventually constrain the program's sustainability.

The overall achievement of 84.60% is comparable with other CIPP evaluations of Indonesian Teaching Factory programs. Irfan et al. (2024), for example, reported Context, Input, Process, and Product percentages of 85.4%, 83.7%, 86.3%, and 83.4% in two agribusiness vocational schools. Yoto et al. (2024) and Saputro et al. (2023) likewise demonstrate how CIPP can distinguish between the formal readiness of a Teaching Factory and the operational quality of its implementation. The convergence does not imply that programs are interchangeable, but it suggests that successful Teaching Factories typically depend on alignment across all four components rather than a single high-performing activity.

Copyright © 2026 The Author(s)

Interpreting the Uniformly High Scores

The near-uniformly favorable ratings deserve methodological scrutiny. All component means lay between 4.19 and 4.28, and even the lowest indicators remained at 4.00 or above. Genuine program strengths may explain part of this pattern, particularly the visible safety routines and industry-linked practice. Nevertheless, 76 of the 81 questionnaire respondents were students, teachers, or managers within the same school, while the five industry respondents were existing program partners. The source instruments were also largely positively worded and requested identifying information such as initials or names. Institutional loyalty, perceived evaluation stakes, acquiescence, courtesy toward the school, and reluctance to criticize a program with which respondents were directly involved could therefore have shifted answers upward.

The purposive eligibility criterion may also contribute to a favorable profile because it deliberately included students who had completed at least one year of full program exposure and did not represent new entrants or students with limited participation. Triangulation reduces, but does not eliminate, this concern: the documents were school records and the implementation ratings came from only two observers. Importantly, the observation percentage was lower than the questionnaire percentage and the four administrative practices rated at 3 show that the evidence was not uniformly positive. The safest interpretation is therefore comparative: the ratings identify stronger and weaker program functions, but the absolute labels 'good' and 'very good' may overstate precision and should not be treated as bias-free measures of performance.

The Context findings are particularly important. Industry partnership goals and policies were the strongest context indicator, and industry support was the strongest input. These results imply that the program's credibility rests heavily on the relationship between the school and automotive-industry partners. This is consistent with research showing that Teaching Factory learning gains value when industrial actors contribute standards, current technical expectations, work culture, and feedback rather than functioning only as occasional guest participants (Ampera et al., 2023; Chasovy et al., 2024). Partnership quality can therefore be understood as a mechanism connecting context with input and process.

The relatively lower score for goal clarity deserves attention even though it remained positive. A Teaching Factory can have strong external partnerships while still producing different internal understandings of its educational and service objectives. Students may see it primarily as practical training, teachers as a competency-assessment environment, and managers as a BLUD service unit. Explicitly articulating shared objectives, performance indicators, and stakeholder responsibilities would reduce this ambiguity and provide a more stable basis for monitoring.

BLUD Flexibility Must Be Matched by Administrative Capability

BLUD status is intended to make public services more responsive, but flexibility does not remove the need for control. The Input profile shows this tension: industry partnership and curriculum readiness were strong, whereas administrative and financial readiness was the lowest Input indicator (M = 4.00). At the same time, operational financial management within Process was rated more highly (M = 4.29). The coexistence of these scores suggests that centralized compliance can function reasonably well at management level while broader administrative capacity and its integration into student learning remain less mature. In other words, adult BLUD personnel may keep transactions orderly without students routinely experiencing the full administrative workflow. The observation scores of 3 for student administration, job recording, simple finance, and service procedures provide direct support for this interpretation.

The implication is pedagogical as well as managerial. Motorcycle service work includes receiving orders, estimating materials and labor, recording parts, explaining results to customers, documenting completion, and reconciling transactions. When financial accountability is concentrated among authorized adults, separating it from students is understandable; however, that separation leaves students exposed mainly to the technical segment of the service chain. The weaker administrative result is therefore plausibly a role-design problem, not simply a shortage of forms. Students need bounded, supervised responsibilities such as opening work orders, recording parts, calculating simple estimates, documenting quality checks, and closing service records. Embedding these responsibilities in job sheets and rotating assignments would preserve formal financial controls while making administration part of occupational competence.

Facilities also require planned renewal. A mean of 4.09 indicates adequacy rather than failure, yet motorcycle technology, diagnostic systems, and service procedures change continuously. BLUD flexibility could be used to establish a transparent reinvestment mechanism in which a defined share of service revenue supports consumables, calibration, preventive maintenance, and priority equipment replacement. Such a mechanism would connect customer-service activity with the long-term quality of the learning environment.

Why Learning Evaluation and Service Quality Lagged

Learning management and evaluation received M = 4.04, markedly below SOP and occupational safety (M = 4.55). The difference is plausible because safety practices are visible, immediate, and readily standardized: observers can see protective behavior, tool control, and procedural compliance during a service task. High-quality learning evaluation requires a more demanding feedback loop, including competency-based rubrics, observation of each student's contribution, timely feedback, periodic review, and documented use of results to adjust teaching. Authentic customer jobs also vary in difficulty and timing, so pressure to complete the service may displace individual assessment and reflection. The present data cannot determine which mechanism was dominant, but the score pattern indicates that procedural compliance was more consistently established than the assessment cycle.

Service or production quality also received M = 4.04. The photographs and the production-management score (M = 4.30) show that real service activity occurred, but activity volume and workflow are not equivalent to verified quality. The study did not report customer-satisfaction ratings, completion time, repeat-service, rework or defect rates, warranty returns, or independent post-service inspection. Moreover, the observed weaknesses in customer-facing and job-record administration can interrupt the information loop through which complaints, defects, and corrective actions are captured. A stronger system would connect before-and-after diagnostic checklists, instructor sign-off, customer feedback, and rework logs to both service-quality assurance and student assessment.

Safety, Authentic Work, and Perceived Readiness

The strongest process result was the application of SOP and occupational health and safety (M = 4.55). This finding is educationally significant because safety routines are not merely compliance requirements; they structure how students plan, execute, and check technical work. Consistent tool use, work-area control, personal protective practices, and service procedures help convert workshop activity into professional practice. The combination of strong safety ratings and field photographs of authentic servicing suggests that the program succeeded in making industrial routines visible within school learning.

The high Product ratings for hard skills, soft skills, and work readiness are compatible with earlier studies that connect Teaching Factory participation to occupational readiness and employability-related outcomes (Khoiron, 2016; Dewi & Sudira, 2018; Prianto et al., 2021).

Copyright © 2026 The Author(s)

Rosidah and Sutirman (2023) similarly emphasize the added value of service-production learning for graduate readiness, while Imran et al. (2024) report relationships among Teaching Factory implementation, work readiness, and future job perspectives. Authentic work can plausibly support these outcomes because students encounter technical uncertainty, customer expectations, time constraints, teamwork, and quality standards in one learning environment.

Crucially, the work-readiness mean of 4.30 is not a pretest-posttest gain score. It is a contemporaneous agreement rating collected from stakeholders at one point in time. Because the student frame intentionally comprised learners with at least one year of full participation, the score may reflect program exposure, but it may also reflect prior motivation, maturation, industrial work-placement experience, socioeconomic conditions, or other unmeasured influences. The single-site design provides no counterfactual showing what readiness would have been without the Teaching Factory or BLUD arrangement. The appropriate conclusion is therefore that stakeholders perceived strong work readiness among program-experienced students in a favorably implemented setting. Demonstrating that the program caused an increase would require baseline and follow-up performance measures, a credible comparison condition, or longitudinal graduate and employer evidence.

Improvement Priorities

The results support continuation of the program, but continuation should be tied to measurable improvement. First, the school should formalize a minimum BLUD-Teaching Factory administrative package: work-order forms, cost estimates, parts and inventory records, service-quality checks, revenue and expenditure records, and periodic management reports. Student learning roles should be embedded in this workflow so that administration becomes part of occupational competence rather than a separate office activity.

Second, monitoring should combine implementation indicators with outcome indicators. Existing questionnaires can track stakeholder experience, but they should be supplemented by rubric-scored technical performance, customer satisfaction, repeat-service rates, completion time, rework or defect rates, certification results, and graduate destinations. A small longitudinal dashboard could follow graduates at six and twelve months and obtain structured employer feedback. This would help distinguish perceived readiness from demonstrated transition outcomes.

Third, industry partnership should move from general support to an annual co-review cycle. Industry representatives and teachers could jointly review curriculum priorities, equipment gaps, job-sheet relevance, safety requirements, service-quality criteria, and teacher upskilling needs. The strong partnership score indicates that the relational foundation already exists; the next step is to convert that strength into documented, recurring quality assurance.

Fourth, facility renewal should be prioritized through risk and curricular relevance. Equipment essential to safety, diagnosis, and frequently performed service jobs should receive first priority. The school can then connect the procurement plan with BLUD revenue, public funding, and industry contributions. This approach avoids treating facility improvement as an occasional project and instead makes it part of program governance.

Limitations and Future Research

Several limitations bound the interpretation. The study involved one Motorcycle Engineering program, and only five industry respondents, which limits transferability and the range of employer perspectives. Its cross-sectional design cannot establish change or causality. Most outcomes were self- or stakeholder-reported; 76 of 81 questionnaire respondents were school insiders, the industry respondents were existing partners, and the source forms requested identifying information and used largely favorable wording. These conditions create plausible social-desirability, acquiescence, courtesy, and ceiling effects. Purposive eligibility followed by

saturated coverage accurately described experienced participants at this site, but it excluded less-exposed students and does not support population-wide generalization. Two observers and school-generated documents strengthened the evidence base but were not fully independent outcome measures. Finally, the source analysis did not report respondent-group distributions, component-level dispersion, confidence intervals, a quantitative content-validity index, post-revision item coefficients, or reliability for each role-specific questionnaire.

Future research should compare multiple BLUD and non-BLUD schools, report results separately by respondent group, use anonymous questionnaires with balanced item wording, and include independent observers. Direct technical-performance rubrics, customer and service-quality measures, certification outcomes, and repeated measurements would help determine whether administrative strengthening, equipment renewal, or expanded industry participation changes student performance. Graduate tracer studies and structured employer assessments are particularly important for testing whether perceived school-based readiness translates into employment, retention, and career development.

Conclusion

The BLUD-based Motorcycle Engineering Teaching Factory at SMK Negeri 2 Tanjungpandan demonstrated a favorable overall CIPP perception profile (M = 4.23; 84.60%). Its central strengths were industry-oriented partnerships, curriculum relevance, consistent SOP and occupational-safety practices, authentic service learning, and highly rated student and learning outcomes. Observation (81.91%) and documentation (85.00%) broadly converged with the questionnaire result while also identifying weaker student involvement in BLUD administration, job recording, simple finance, and service procedures. Priority needs concern administrative and financial learning integration, facility renewal, competency-based learning evaluation, service-quality assurance, and systematic monitoring. The program is suitable for continuation with targeted strengthening, but the uniformly high ratings should be interpreted cautiously because of possible response and ceiling bias. Most importantly, the cross-sectional, single-site evidence supports strong implementation and perceived readiness among experienced participants; it does not show that the Teaching Factory or BLUD arrangement caused an increase in work readiness. Future evaluation should connect the CIPP profile with anonymous multi-stakeholder measures, direct performance evidence, customer outcomes, and longitudinal graduate and employer data.

References

Ampera, D., Tanjung, N., Farihah, F., Sinukaban, V. Y., & Juliana, N. (2023). Model of teaching factory in vocational high school fashion program students. *Jurnal Pendidikan Teknologi dan Kejuruan, 29*(2), 189–207. https://doi.org/10.21831/jptk.v29i2.65043

Chasovy, A., Jalinus, N., Waskito, W., & Wulansari, R. E. (2024). Analysis of the implementation of the Teaching Factory program in vocational high school. *Jurnal Penelitian Pendidikan IPA, 10*(7), 4364–4375. https://doi.org/10.29303/jppipa.v10i7.7569

Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications.

Dewi, S. S., & Sudira, P. (2018). The contribution of Teaching Factory program implementation on work readiness of vocational high school students in Makassar. *Journal of Educational Science and Technology, 4*(2), 126–131. https://doi.org/10.26858/est.v4i2.6434

Copyright © 2026 The Author(s)

Diwangkoro, E., & Soenarto, S. (2020). Development of Teaching Factory learning models in vocational schools. *Journal of Physics: Conference Series, 1456*, 012046. https://doi.org/10.1088/1742-6596/1456/1/012046

Frye, A. W., & Hemmer, P. A. (2012). Program evaluation models and related theories: AMEE Guide No. 67. *Medical Teacher, 34*(5), e288–e299. https://doi.org/10.3109/0142159X.2012.668637

Hariyanto, W. (2021). Desain implementasi Badan Layanan Umum Daerah pada sekolah menengah kejuruan sebagai penguatan pembelajaran berbasis industri. *JAMP: Jurnal Administrasi dan Manajemen Pendidikan, 4*(2), 121–138. https://doi.org/10.17977/um027v4i12021p121

Imran, I., Marji, M., Suswanto, H., & Adhikari, B. P. (2024). The influence of Teaching Factory (TEFA) implementation and work readiness on vocational high school students' future job perspectives. *Jurnal Pendidikan Vokasi, 14*(1), 86–96. https://doi.org/10.21831/jpv.v14i1.66796

Irfan, A. M., Romadin, A., Putra, I. I., Baharsyah, N. I., & Nawi, M. Z. M. (2024). Evaluating the effectiveness of the Teaching Factory learning model in leading vocational schools: A case study of agribusiness, food crops, and horticulture programs. *Journal of Vocational Education Studies, 7*(2), 227–246. https://doi.org/10.12928/joves.v7i2.10485

Kautsar, A., Wiyono, G., Mulia, M., Iqbal, M., & Al-Fairusy, M. (2022). Teaching Factory model development in vocational high schools. *AL-ISHLAH: Jurnal Pendidikan, 14*(4), 6347–6360. https://doi.org/10.35445/alishlah.v14i4.2461

Khoiron, A. M. (2016). The influence of Teaching Factory learning model implementation to the students' occupational readiness. *Jurnal Pendidikan Teknologi dan Kejuruan, 23*(2), 122–129. https://doi.org/10.21831/jptk.v23i2.12294

Ministry of Home Affairs of the Republic of Indonesia. (2018). Regulation of the Minister of Home Affairs Number 79 of 2018 concerning Regional Public Service Agencies.

Nouraey, P., Al-Badi, A., Riasati, M. J., & Maata, R. L. (2020). Educational program and curriculum evaluation models: A mini systematic review of the recent trends. *Universal Journal of Educational Research, 8*(9), 4048–4055. https://doi.org/10.13189/ujer.2020.080930

Prianto, A., Winardi, W., & Qomariyah, U. N. (2021). The effect of the implementation of Teaching Factory and its learning involvement toward work readiness of vocational school graduates. *International Journal of Instruction, 14*(1), 283–302. https://doi.org/10.29333/iji.2021.14117a

Ratnaya, G., Indriaswuri, R., Widayanthi, D. G. C., Atmaja, I. M. P. D., & Dalem, A. A. G. P. K. P. (2022). CIPP evaluation model for vocational education: A critical review. *Education Quarterly Reviews, 5*(3), 1–8. https://doi.org/10.31014/aior.1993.05.03.519

Rosidah, R., & Sutirman, S. (2023). Added value of Teaching Factory learning in services production unit to prepare graduate work readiness. *Cakrawala Pendidikan, 42*(3), 695–704. https://doi.org/10.21831/cp.v42i3.49137

Saputro, I. N., Nurhidayati, A., Suharno, S., Winiawati, C., Yunika, R., & Kusumadevi, M. P. (2023). Evaluation of the Factory implementation of the CIPP method in Indonesian vocational education. *AL-ISHLAH: Jurnal Pendidikan, 15*(4), 4972–4982. https://doi.org/10.35445/alishlah.v15i4.3903

Stufflebeam, D. L., & Shinkfield, A. J. (2007). *Evaluation theory, models, and applications*. Jossey-Bass.

Suharno, S., Pambudi, N. A., & Harjanto, B. (2020). Vocational education in Indonesia: History, development, opportunities, and challenges. *Children and Youth Services Review, 115*, 105092. https://doi.org/10.1016/j.childyouth.2020.105092

Wibawa, E. A., Triyono, M. B., Setiadi, B. R., & Pratama, G. N. I. P. (2023). Empirical study of management of vocational high school with Regional Public Services Agency in Indonesia. *WSEAS Transactions on Business and Economics, 20*, 52–60. https://doi.org/10.37394/23207.2023.20.6

Yoto, Y., Marsono, M., Qolik, A., & Romadin, A. (2024). Evaluation of Teaching Factory using CIPP (Context, Input, Process, Product) model to improve vocational high school students' skills. *Jurnal Pendidikan Vokasi, 14*(1), 12–28. https://doi.org/10.21831/jpv.v14i1.62573