JPPIPA 11(10) (2025)

Jurnal Penelitian Pendidikan IPA

http://jppipa.unram.ac.id/index.php/jppipa/index

# Evaluating the Impact of Self-Directed PMM Digital Training Platform on Teacher Professional Competencies Using the Kirkpatrick Evaluation Model

Suharry[1], Fahmi Rizal[2*], M. Giatman[2], Syaiful Haq[2], Lasyatta Syaifullah[1]

[1] Technology and Vocational Education, Universitas Negeri Padang, Padang, Indonesia.
[2] Departement of Civil Engineering, Universitas Negeri Padang, Padang, Indonesia.

Received: July 19, 2025
Revised: September 08, 2025
Accepted: October 25, 2025
Published: October 31, 2025

Corresponding Author:
Fahmi Rizal
fahmi@ft.unp.ac.id

DOI: 10.29303/jppipa.v11i10.12228



**Abstract:** Civil servant teachers at SMK Negeri 2 Sungai Penuh face challenges in applying Platform Merdeka Mengajar (PMM) self-directed training results to classroom practices. This study evaluated the effectiveness of the PMM self-directed training program using the Kirkpatrick evaluation model. An evaluative research design was employed involving 20 civil servant teachers selected through purposive sampling. Data were collected using questionnaires, interviews, observations, tests, and documentation, then analyzed using SPSS Statistics Version 29 for descriptive and inferential statistics (paired t-test, effect size), and thematic analysis. Results demonstrated program effectiveness across all four Kirkpatrick levels. At the reaction level, teachers showed high satisfaction (M = 3.65-4.20). The learning level revealed significant competency improvements with paired t-test results ($t = -24.367$, $p < 0.001$) and exceptionally large effect size (Cohen's $d = -5.449$). At the behavior level, positive teaching behavior changes emerged (M = 3.85-4.00) across four competency dimensions. The results level identified transformative impacts including improved student learning outcomes, innovative methodologies (flipped classroom, gamification), efficient resource utilization, and optimized curriculum achievement. In conclusion, the PMM self-directed training program effectively improved teacher competency and transformed teaching practices at SMK Negeri 2 Sungai Penuh.

**Keywords:** Platform Merdeka Mengajar; Kirkpatrick Evaluation Model; Self-Directed Training; Teacher Professional Development; Vocational Education.

## Introduction

The rapid advancement of digital technology and the demands of Industry 4.0 have fundamentally transformed educational paradigms worldwide, necessitating a shift from traditional teacher-centered approaches to technology-enhanced, student-centered learning environments (Lapisa et al., 2025). Vocational education, as defined by Law Number 20 of 2003 concerning the National Education System, serves a strategic role in producing skilled and competitive human resources aligned with labor market needs by preparing students to work in specific fields (Republic of Indonesia, 2003). In this context, teacher professional development has become a critical determinant of educational quality, as teachers must continuously adapt their competencies to meet the evolving demands of digital-age pedagogy and industrial workplace requirements.

In Indonesia, the Merdeka Belajar policy framework mandates the transformation of teacher competencies through Director General of Teachers and

**How to Cite:**
Suharry, Rizal, F., Giatman, M., Haq, S., & Syaifullah, L. (2025). Evaluating the Impact of Self-Directed PMM Digital Training Platform on Teacher Professional Competencies Using the Kirkpatrick Evaluation Model. *Jurnal Penelitian Pendidikan IPA*, *11*(10), 495–504. https://doi.org/10.29303/jppipa.v11i10.12228

Education Personnel Regulation No. 2626/B/HK.04.01/2023, which establishes updated professional standards for educators. According to Law No. 14 of 2005 concerning Teachers and Lecturers, teacher competence encompasses four interrelated dimensions: pedagogical, personal, social, and professional competence (Republic of Indonesia, 2005). Pedagogical competency encompasses teachers' ability to manage learning, understand student characteristics, design curricula, conduct learning outcome evaluations, and develop students' potential holistically (Mulyasa, 2021). Personal competence reflects teachers' personal qualities as role models, including maturity, morality, work ethic, and professional integrity (Uno, 2020). Social competence involves effective communication and interaction with education stakeholders (Suprihatiningrum, 2022), while professional competence emphasizes mastery of scientific substance, structure, and teaching methods coherent with subject matter (Sagala, 2019). To operationalize these competency requirements, the Ministry of Education, Culture, Research, and Technology developed the Platform Merdeka Mengajar (PMM), a digital innovation supporting continuous teacher professional development through self-directed training centered on student learning needs (Kemdikbudristek, 2022).

Despite the widespread implementation of PMM self-directed training programs across Indonesian vocational schools, a critical gap exists in systematic evaluation practices to assess program effectiveness. Training modules covering learner-centered pedagogy, learning assessment implementation, and innovative teaching methodologies have been deployed on a large scale, yet empirical evidence regarding their impact on teacher competency development and classroom practice transformation remains limited. At SMK Negeri 2 Sungai Penuh, civil servant teachers face specific challenges in applying PMM training outcomes, particularly in implementing student-centered learning strategies, optimizing inter-teacher collaboration, managing effective learning processes, and mastering content-pedagogical integration aligned with professional competency standards. The absence of comprehensive evaluation creates uncertainty regarding program success in achieving intended learning outcomes, behavior change, and organizational impact. This evaluation deficit poses serious implications for evidence-based policy development, resource allocation optimization, and continuous improvement of digital teacher training initiatives. Furthermore, without rigorous assessment frameworks, stakeholders cannot determine whether substantial investments in digital training platforms yield measurable returns in terms of enhanced teaching quality and improved student learning outcomes.

To address this issue, this research employs the Kirkpatrick evaluation model, a gold-standard framework for training program assessment introduced in 1959 and widely validated across educational contexts (Kirkpatrick & Kirkpatrick, 2016). The model's four-level structure—reaction, learning, behavior, and results—provides a systematic approach to evaluating training effectiveness while considering factors influencing learning transfer to workplace practice (McLaughlin & Jordan, 2015). Recent empirical studies have confirmed the model's reliability and validity in vocational and educational training contexts. Hidayat et al. (2023) successfully applied the Kirkpatrick framework to evaluate an EYD V training program at SMK Broadcasting Mahardika, demonstrating positive participant responses, improved competencies, behavioral changes, and enhanced student outcomes. Romdon & Arief (2024) found similar effectiveness patterns in evaluating a P5 Entrepreneurship program at SMK PGRI 3 Bogor, with high satisfaction, increased knowledge, attitude changes, and improved student engagement. In international contexts, Makinde & Bamiro (2024) evaluated a TVET program in Nigeria with 450 participants, confirming program satisfaction, technical competency improvements, workplace skill application, and organizational productivity gains.

The robustness of the Kirkpatrick model is further supported by diverse training evaluation studies. Farida & Nurhadi (2023) evaluated industrial-standard GTAW welding training, finding positive responses, improved welding skills, safety standard application, and enhanced outcome quality. Alsalamah & Callinan (2021) adapted the model for principal training evaluation, yielding high satisfaction, improved management knowledge, effective leadership implementation, and enhanced school performance. Mahmoodi et al. (2019) assessed an in-service teacher training program in Iran, identifying satisfaction with content, improved pedagogical competencies, teaching method changes, and better student learning outcomes. Ramadhan et al. (2022) applied the framework to BIM training evaluation, documenting material satisfaction, improved BIM understanding, project application, and work efficiency gains. Du (2021) examined the training evaluation system at Zhejiang Open University, revealing learning system satisfaction, improved teaching competencies, learning method innovation, and education quality enhancement. Imansari et al. (2023) evaluated PLC training implementation, finding positive responses, improved PLC knowledge, programming application, and enhanced technical skills.

However, existing literature reveals a significant research gap: no prior study has comprehensively evaluated digital self-directed teacher training platforms

in Indonesian vocational education contexts using the complete four-level Kirkpatrick framework with mixed-methods analysis. While previous research has examined individual aspects of training effectiveness, none has integrated quantitative competency assessment (pre-post testing with effect size analysis), behavioral observation across multiple competency dimensions, and qualitative analysis of organizational impact within a single systematic evaluation. This study addresses this gap by providing the first comprehensive Kirkpatrick-based evaluation of the PMM self-directed training program, specifically examining its effectiveness in transforming teacher competencies and teaching practices in vocational school contexts.

The urgency of this research stems from four critical imperatives. First, without systematic evaluation evidence, policymakers lack empirical foundations for scaling or modifying the PMM platform, risking inefficient resource allocation in national teacher development initiatives. Second, vocational schools require validated assessment frameworks to demonstrate accountability and justify continued investment in digital training infrastructure. Third, empirical data on training effectiveness is essential for aligning teacher professional development with national competency standards outlined in Director General of GTK Regulation No. 2626/B/HK.04.01/2023. Fourth, identifying effective components and implementation barriers within PMM training can inform evidence-based improvements to maximize program impact on teaching quality and student learning outcomes.

Therefore, this study aims to conduct a comprehensive evaluation of the PMM self-directed training program at SMK Negeri 2 Sungai Penuh using the Kirkpatrick four-level evaluation model. Specifically, the research objectives are to: (1) evaluate teachers' reaction levels to the PMM self-directed training program in terms of satisfaction, relevance, and perceived utility; (2) analyze learning levels achieved by teachers through pre-test and post-test competency assessments with effect size calculations; (3) examine behavioral changes in teaching practice across four competency dimensions—pedagogical, personal, social, and professional—as specified in Director General of Teacher and Education Personnel Regulation No. 2626/B/HK.04.01/2023; and (4) analyze program results regarding impacts on classroom teaching effectiveness, student learning outcomes, innovative teaching methodology adoption, and inter-teacher collaboration in instructional development. Through this systematic mixed-methods evaluation, the study provides empirical evidence of PMM training effectiveness and offers actionable recommendations for enhancing future digital teacher professional development programs in vocational education contexts.

## Method

This research is an evaluative study employing the Kirkpatrick four-level evaluation model to assess the effectiveness of the Platform Merdeka Mengajar (PMM) self-directed training program. The research adopts a mixed-methods approach, combining quantitative and qualitative data collection and analysis techniques to provide comprehensive evaluation across four levels: reaction, learning, behavior, and results. The research was conducted from August to October 2024 at SMK Negeri 2 Sungai Penuh, Jambi Province, Indonesia. The school was selected based on its active implementation of the PMM self-directed training program and the availability of civil servant teachers who had completed the training modules. The research population consisted of 42 civil servant (ASN) teachers at SMK Negeri 2 Sungai Penuh. Using purposive sampling technique, 20 civil servant teachers were selected as research respondents based on the following criteria: (1) active civil servant status at the school, (2) completion of at least three PMM self-directed training modules, (3) minimum of two years teaching experience, and (4) willingness to participate in all stages of the research. The sample size was determined based on the adequacy for parametric statistical analysis and the availability of teachers meeting the inclusion criteria.

The research was conducted through six systematic stages:

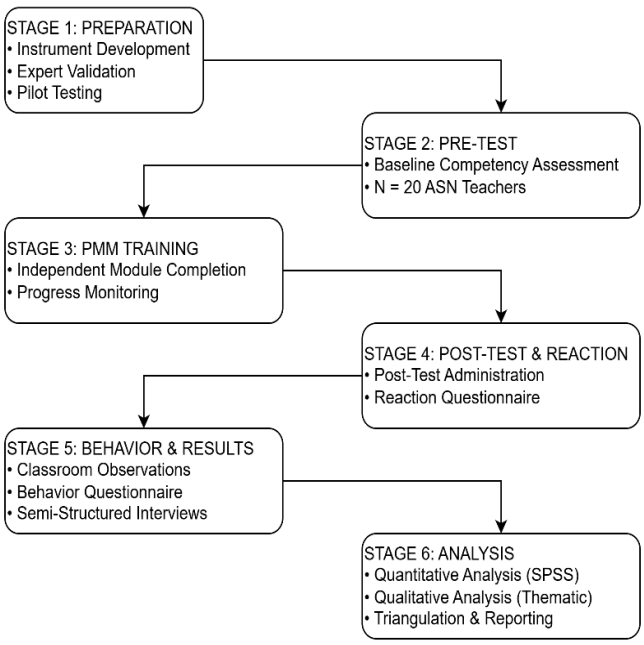


**Figure 1.** Flowchart of the Research

Stage 1: Preparation and Instrument Development, including questionnaires, test instruments, interview guidelines, and observation sheets. Instruments were validated through expert judgment and pilot testing. Stage 2: Pre-Test Administration, Pre-tests were administered to all 20 participating teachers to measure baseline competency levels before they engaged with PMM training modules. Tests were conducted in controlled conditions to ensure data validity. Stage 3: PMM Training Implementation, Teachers independently completed selected PMM training modules on the platform, including modules on student-centered learning, learning assessment, and innovative teaching methods. Training progress was monitored through platform analytics. Stage 4: Post-Test and Reaction Level Data Collection, Post-tests were administered to measure competency improvements following training completion. Simultaneously, reaction level questionnaires were distributed to assess teachers' immediate responses to the training program. Stage 5: Behavior Level and Results Level Data Collection, Classroom observations were conducted to document behavior changes in teaching practices. Behavior level questionnaires were distributed to principals and peer teachers for triangulation. Semi-structured interviews were conducted with participating teachers to gather qualitative data on program results and impacts. Stage 6: Data Analysis and Reporting, Comprehensive data analysis was performed using multiple techniques, followed by interpretation, triangulation, and report preparation. The data collection instruments included questionnaires, interview guidelines, observation sheets, test instrument sheets, and documentation study sheets. Data analysis techniques included, quantitative descriptive analysis using SPSS Statistics Version 29 software, parametric inferential statistics for pre-test and post-test difference tests, and thematic analysis for qualitative data at the results level.

## Result and Discussion

The results of this study are presented systematically according to the four levels of Kirkpatrick's evaluation model: reaction, learning, behavior, and results. Each evaluation level provides in-depth insights into the effectiveness of the training program in achieving the goals of developing civil servant teachers' competencies.

*Learning Level Results*

The data analysis at the reaction level was conducted using quantitative descriptive analysis techniques with SPSS Statistics version 29 software to process questionnaire data. The analytical criteria included calculations of mean, median, mode, standard deviation, and percentages to determine respondents' reaction levels to the Independent Learning Platform Merdeka Mengajar (PMM) training program. The results of the analysis of respondent reaction levels are shown in Table 1.

**Table 1.** Frequency Distribution of Reaction Level Data

| | | *Statistics* | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Digital Content Quality | | | | Platform Accessibility | | | | Implementation Flexibility | | | |
| | | A1 | A2 | A3 | A4 | B1 | B2 | B3 | B4 | C1 | C2 | C3 | C4 |
| N | Valid | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 |
| | Missing | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Mean | | 4.10 | 4.15 | 4.20 | 3.70 | 4.05 | 3.75 | 3.65 | 4.00 | 4.10 | 4.15 | 3.90 | 3.95 |
| Median | | 4.00 | 4.00 | 4.00 | 4.00 | 4.00 | 4.00 | 4.00 | 4.00 | 4.00 | 4.00 | 4.00 | 4.00 |
| Std. Deviation | | .641 | .366 | .410 | .470 | .510 | .550 | .489 | .459 | .308 | .366 | .447 | .394 |
| Minimum | | 3 | 4 | 4 | 3 | 3 | 3 | 3 | 3 | 4 | 4 | 3 | 3 |
| Maximum | | 5 | 5 | 5 | 4 | 5 | 5 | 4 | 5 | 5 | 5 | 5 | 5 |

The frequency distribution results of reaction level data presented in Table 1 show that teachers' responses to the training program were very positive, with mean values ranging from 3.65 to 4.20 across all measured indicators. The highest value was obtained in the digital content quality aspect (A3=4.20), followed by implementation flexibility (C2=4.15) and platform accessibility (B1=4.05). The consistent median value of 4.00 across all indicators indicates that the majority of participants provided good to very good ratings of the training program. The relatively low standard deviation values (0.308-0.641) demonstrate good response homogeneity among participants, while the range of minimum values of 3 to maximum values of 5 confirms that all respondents provided ratings in the good to very good categories without any negative assessments of the training program.

*Hasil Level Learning*

The data analysis at the learning level was conducted using parametric inferential statistics consisting of prerequisite analysis tests including normality and homogeneity tests, as well as paired samples t-test and effect size calculations using SPSS Statistics version 29 software. The normality test used

the Shapiro-Wilk test, and the homogeneity test used Levene's test. Subsequently, to test significant differences between pre-test and post-test scores, a paired samples t-test was used, and to measure the magnitude of effect of the training program, effect size was calculated using Cohen's d and Hedges' correction. The normality test results are presented in Table 2.

**Table 2.** Normality Test Results

| | Kolmogorov-Smirnov[a] | | | Shapiro-Wilk | | |
|---|---|---|---|---|---|---|
| | Statistic | df | Sig. | Statistic | df | Sig. |
| Nilai *Pre-Test* | .107 | 20 | .200* | .980 | 20 | .929 |
| Nilai *Post-Test* | .168 | 20 | .141 | .971 | 20 | .766 |

*. This is a lower bound of the true significance.
a. Lilliefors Significance Correction

The normality test results presented in Table 2 show that pre-test and post-test value data are normally distributed with Shapiro-Wilk significance values of 0.929 for pre-test and 0.766 for post-test ($p > 0.05$). This indicates that the data meet the normality assumptions required for parametric analysis. After obtaining the normality test results, homogeneity analysis was subsequently conducted with results as shown in Table 3.

**Table 3.** Homogeneity Test Results

| | | Levene Statistic | df1 | df2 | Sig. |
|---|---|---|---|---|---|
| Nilai | Based on Mean | 2.248 | 1 | 38 | .142 |
| | Based on Median | 2.137 | 1 | 38 | .152 |
| | Based on Median and with adjusted df | 2.137 | 1 | 36.791 | .152 |
| | Based on trimmed mean | 2.248 | 1 | 38 | .142 |

The homogeneity test results presented in Table 3 show that the variances of pre-test and post-test data are homogeneous with a Levene's test significance value of 0.142 ($p > 0.05$). These results confirm that the data meet the homogeneity of variances assumption required for conducting paired samples t-test. After meeting the prerequisite analysis tests, a paired samples t-test was subsequently conducted with results as shown in Table 4.

**Table 4.** Paired Samples T-Test Results

| | | Paired Differences | | | | | t | df | Significance | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Mean | Std. Deviation | Std. Error Mean | 95% Confidence Interval of the Difference | | | | One-Sided p | Two-Sided p |
| | | | | | Lower | Upper | | | | |
| Pair 1 | Nilai *Pre-Test* - Nilai *Post-Test* | -20.000 | 3.671 | .821 | -21.718 | -18.282 | -24.367 | 19 | <.001 | <.001 |

The paired samples t-test results presented in Table 4 show that there is a highly significant difference between pre-test and post-test values with $t = -24.367$ ($df = 19$) and significance $p < 0.001$. The mean difference of -20.000 points with a 95% confidence interval between -21.718 and -18.282 indicates a very substantial competency improvement after participating in the training program. Subsequently, to determine the practical impact magnitude of the program on participants' knowledge and skills improvement, effect size analysis results were obtained as shown in Table 5.

**Table 5.** Effect Size Test Results

| | | | Standardizer[a] | Point Estimate | 95% Confidence Interval | |
|---|---|---|---|---|---|---|
| | | | | | Lower | Upper |
| Pair 1 | Nilai *Pre-Test* - Nilai *Post-Test* | Cohen's d | 3.671 | -5.449 | -7.216 | -3.671 |
| | | Hedges' correction | 3.824 | -5.230 | -6.927 | -3.524 |

a. The denominator used in estimating the effect sizes.
Cohen's d uses the sample standard deviation of the mean difference.
Hedges' correction uses the sample standard deviation of the mean difference, plus a correction factor.

The effect size test results presented in Table 5 show that the training program has a very large practical impact with Cohen's d value of -5.449 (95% CI: -7.216 to -3.671) and Hedges' correction of -5.230 (95% CI: -6.927 to -3.524). These very high effect size values indicate that the training program not only produces statistically significant improvements but also has an extraordinarily large practical impact in enhancing participants' competencies, exceeding the standard criteria for very large effects (large effect > 0.8).

*Behavior Level Results*

The data analysis at the behavior level was conducted using quantitative descriptive analysis techniques with SPSS Statistics version 29 software to process questionnaire data. The analytical criteria included calculations of mean, median, mode, standard deviation, and percentages to determine respondents' reaction levels. The analysis results of changes in teachers' classroom teaching behavior after participating in the PMM independent training program are shown in Table 6.

**Table 6.** Frequency Distribution of Behavior Level Data

| | | Implementation of Training Results | | | | Changes in Teaching Behavior | | | | Learning Implementation | | | | Learning Management | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | A1 | A2 | A3 | A4 | B1 | B2 | B3 | B4 | C1 | C2 | C3 | C4 | D1 | D2 | D3 | D4 |
| N | Valid | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 |
| | Missing | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Mean | | 3.90 | 4.00 | 3.95 | 3.95 | 3.90 | 3.85 | 3.90 | 3.90 | 3.90 | 3.95 | 3.90 | 3.85 | 3.85 | 3.90 | 3.90 | 3.85 |
| Median | | 4.00 | 4.00 | 4.00 | 4.00 | 4.00 | 4.00 | 4.00 | 4.00 | 4.00 | 4.00 | 4.00 | 4.00 | 4.00 | 4.00 | 4.00 | 4.00 |
| Std. Deviation | | .447 | .562 | .394 | .510 | .447 | .366 | .447 | .447 | .447 | .394 | .447 | .366 | .366 | .447 | .447 | .366 |
| Minimum | | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 |
| Maximum | | 5 | 5 | 5 | 5 | 5 | 4 | 5 | 5 | 5 | 5 | 5 | 4 | 4 | 5 | 5 | 4 |

The behavior level data analysis results presented in Table 6 show that positive changes in teaching behavior occurred across all measured dimensions. Mean values range from 3.85 to 4.00 with an overall average of 3.90, indicating that teachers successfully implemented training results into classroom teaching practice. In the training implementation dimension, mean values range from 3.90 to 4.00 with the highest value on indicator A2 (4.00). The teaching behavior change dimension shows consistent mean values of 3.85-3.90, indicating good adaptation to new learning approaches. The learning implementation dimension has mean values of 3.85-3.95 with the highest value on indicator C2 (3.95), while the learning management dimension shows mean values of 3.85-3.90. The consistent median value of 4.00 across all indicators shows that the majority of teachers provide good ratings for learning transfer to classroom practice. The relatively low standard deviation values (0.366-0.562) indicate good response homogeneity, with minimum values of 3 to maximum values of 5 showing that all respondents provided ratings in the good to very good categories.

*Results Level Findings*

The data analysis at the results level used thematic analysis techniques to identify patterns and themes related to program impact on learning and program impact on schools. Thematic analysis was conducted through a systematic process including data familiarization, initial coding, theme searching, theme review, defining and naming themes, and compiling analysis reports. The results level data analysis yielded core conclusions from respondent interview answers as shown in Table 7.

**Table 7.** Collection of Core Conclusions from Respondent Interview Answers

| Variable | Questionn | Respondent Answers |
|---|---|---|
| Impact on Learning | 1. How has student learning outcomes improved after implementing PMM training results? | Average respondent answers show significant improvement in student learning outcomes. Improvements are evident in aspects of grades, student learning motivation, active participation in learning, and technology usage skills. |
| | 2. Are there significant changes in student learning achievements? | Average respondent answers confirm significant changes in learning achievements. Students show better learning independence, improved evaluation results, and rapidly developing technology capabilities. |
| | 3. How is the efficiency of learning resource utilization after following PMM? | Average respondent answers show improved efficiency in learning resource utilization. Digital material usage reduces dependence on printed materials, teaching preparation time is more optimal, and learning costs become more economical. |
| | 4. What innovations are there in learning resource utilization? | Average respondent answers identify various implemented learning innovations. Innovations include flipped classroom learning, interactive multimedia, learning gamification, real-time digital assessment, and technology-based learning content development. |
| | 5. To what extent are curriculum targets achieved after PMM implementation? | Average respondent answers show optimal curriculum target achievement. Project-based learning is more easily achieved, formative and summative assessments are more effective, and technology |

| Variable | Questionn | Respondent Answers |
|---|---|---|
| | | integration in learning according to curriculum demands is successfully implemented. |
| | 6. What curriculum target achievement strategies are implemented? | Average respondent answers outline comprehensive applied strategies. Strategies include implementing digital learning models, using interactive multimedia, technology-based assessment, collaborative learning, and digital student progress monitoring. |
| | 7. How is the effectiveness of the learning process after PMM? | Average respondent answers show significant improvement in learning effectiveness. Student participation increases, learning time is more structured, teacher-student interaction is more dynamic, and learning feedback is faster and more accurate. |
| | 8. What are the significant changes in the learning process? | Average respondent answers identify comprehensive learning transformation. Changes include student-centered learning, consistent technology implementation, continuous authentic assessment, soft skills integration, and increased creativity in designing learning. |
| Impact on Schools | 1. How has the culture of professional development changed in schools? | Average respondent answers show positive professional culture transformation. A solid teacher learning community is formed, digital content sharing culture strengthens, active regular discussion forums run, and openness to learning innovation increases. |
| | 2. What is the impact on teacher self-development motivation? | Average respondent answers show significant increase in self-development motivation. Teachers are motivated to continue advanced digital training, actively create original learning content, join digital teacher communities, and participate in ongoing mentoring programs. |
| | 3. How has collaboration among teachers improved after PMM? | Average respondent answers show significant improvement in collaboration. Forms of collaboration include digital learning media development teams, sharing question banks and materials, peer review lesson plans, and collaboration in classroom action research. |
| | 4. What are the benefits of this collaboration? | Average respondent answers identify comprehensive collaboration benefits. Benefits for teachers include improved digital competency, expanded professional networking, and increased teaching motivation. Benefits for schools include standardized learning quality, resource efficiency, and improved school reputation. |
| | 5. How does PMM contribute to achieving school vision and mission? | Average respondent answers show PMM's significant contribution to achieving school vision and mission. Technology integration in learning according to school vision is successfully implemented, quality of graduates ready for work and entrepreneurship improves, and innovative learning according to school vision is achieved. |
| | 6. To what extent does the program support school goals? | Average respondent answers confirm highly significant program support for school goals. The program supports goals of improving learning quality, teacher HR development, learning modernization, and improving school accreditation values. |
| | 7. What is the impact on school learning quality? | Average respondent answers show measurable improvement in learning quality. Learning process quality improves, teacher competency in learning technology develops, student satisfaction with learning increases, and student learning evaluation results show improvement. |
| | 8. What are the supporting factors for this quality improvement? | Average respondent answers identify internal and external supporting factors. Internal factors include principal support, teacher motivation to develop, and adequate IT infrastructure. External factors include Education Office support, stable internet access, and conducive government PMM policies. |

Based on the data presented in Table 7, the PMM independent training program produces significant transformative impacts on two main dimensions: impact on learning and impact on schools. In the learning impact dimension, respondents report significant improvement in student learning outcomes, positive changes in learning achievements, and increased efficiency in learning resource utilization. Implemented learning innovations include flipped classroom, interactive multimedia, learning gamification, and real-time digital assessment. Curriculum target achievement becomes more optimal with comprehensive digital learning strategies, while learning process effectiveness increases through comprehensive transformation

toward student-centered learning. In the school impact dimension, positive transformation of professional development culture occurs with the formation of a solid teacher learning community. Teacher self-development motivation increases significantly, inter-teacher collaboration strengthens through various forms of cooperation, and the program's contribution to achieving school vision and mission is highly significant. School learning quality experiences measurable improvement with support from conducive internal and external factors.

*Discussion*

The main findings of this research demonstrate the effectiveness of the program across all four evaluation levels. At the reaction level, the study shows consistent positive responses from teachers toward the PMM self-directed training program. Descriptive statistical analysis results show mean values above 3.5 for all questionnaire items, with a range of values between 3.65 and 4.20. The digital content quality dimension received the highest rating with a mean value of 4.20, indicating that teachers positively assess the quality of learning materials provided by the platform. The consistency of median values at 4.00 for all items shows that the majority of respondents provided "agree" ratings for various aspects of the training program, including platform accessibility and implementation flexibility. These findings confirm that the training program successfully met teachers' expectations at the initial evaluation stage.

The learning level demonstrates significant impact of the training program on improving teacher competency. Paired samples t-test results reveal a highly significant difference between Pre-Test and Post-Test scores with $t = -24.367$ and significance $p < 0.001$. The mean difference of 20.000 points between pre- and post-training conditions indicates substantial competency improvement. Furthermore, effect size analysis using Cohen's d yielded a value of -5.449, indicating a very large practical impact that exceeds standard criteria for extraordinarily significant effects. These findings prove that the training program not only provides statistical differences but also highly meaningful practical impact in the context of teacher professional development.

At the behavior level, the study identifies positive and consistent changes in teaching behavior following training program implementation. Statistical analysis shows mean values ranging from 3.85 to 4.00 across four evaluated competency dimensions: training implementation, teaching behavior change, learning implementation, and learning management. The consistency of median values at 4.00 for all items indicates that teachers have successfully transferred learning into actual classroom practice. The relatively small standard deviation (0.366-0.562) shows homogeneity in teaching behavior changes among respondents, confirming the program's success in facilitating learning transfer from theory to practice.

The results level reveals transformative impacts of the training program on learning quality and institutional school development. Thematic analysis identifies two main themes encompassing impact on learning and impact on schools. In the learning dimension, eight sub-themes were found showing improved student learning outcomes, significant changes in learning achievements, efficient use of learning resources, and innovation in learning methodology. Teachers reported implementing various methodological breakthroughs such as flipped classroom, interactive multimedia development, learning gamification implementation, and real-time digital assessment. Impact on schools includes optimization of curriculum target achievement, effectiveness of formative-summative assessment, and successful technology integration according to curriculum demands.

The findings of this research show high consistency with previous studies that used the Kirkpatrick evaluation model in the context of teacher training programs and professional development. At the reaction level, the findings align with research by Hidayat et al. (2023) which showed positive participant responses to the EYD V usage training program at SMK Broadcasting Mahardika. Both studies demonstrate that high satisfaction levels at the initial evaluation stage become important indicators of training program success. This consistency of findings is also supported by research by Romdon & Arief (2024) which identified high participant satisfaction levels in the P5 Entrepreneurship program at SMK PGRI 3 Bogor. This similarity in patterns indicates that positive responses at the reaction level are common characteristics of well-designed training programs that meet participant needs.

Findings at the learning level show alignment with research by Makinde & Bamiro (2024) which evaluated TVET programs in Nigeria and found significant improvement in technical competency after training program implementation. The very large effect size value in this study (Cohen's d = -5.449) even exceeds findings from research by Mahmoodi et al. (2019) which evaluated in-service teacher training programs in Iran with significant pedagogical competency improvements. This comparison shows that the PMM self-directed training program has extraordinarily effective learning impact, even when compared to conventional training programs requiring face-to-face interaction.

Results at the behavior level show consistency with findings by Farida & Nurhadi (2023) which identified

implementation of work safety standards and improved welding quality after GTAW welding training. Both studies demonstrate that learning transfer from training conditions to actual practice can be achieved effectively when training programs are designed with appropriate approaches. These findings also align with research by Azmy & Setiarini (2023) which identified implementation of assessment standards in assessor training programs, confirming that behavior change is a critical indicator of professional development program success.

At the results level, research findings show alignment with research by Alsalamah & Callinan (2021) which adapted the Kirkpatrick evaluation model for principal training programs and found significant school performance improvements. Learning methodology innovations found in this study, such as flipped classroom implementation, interactive multimedia development, and learning gamification application, align with findings by Ramadhan et al. (2022) which identified BIM implementation in projects and improved work efficiency. This consistency of findings indicates that technology-based training programs have high transformative potential in changing professional practice.

However, this research also shows several advantages compared to previous studies. The very large effect size value and transformative impact at the results level demonstrate that the PMM self-directed training program has superior effectiveness compared to conventional training programs. Findings by Du (2021) which evaluated training systems at Zhejiang Open University showed positive results but did not achieve the magnitude of effect as large as this research. This difference is likely due to the unique characteristics of the PMM self-directed training program that provides high flexibility, optimal accessibility, and high-quality digital content that suits the contextual needs of vocational school teachers.

## Conclusion

The evaluation of the self-directed training program on the Platform Merdeka Mengajar (PMM) at SMK Negeri 2 Sungai Penuh using the Kirkpatrick evaluation model demonstrates that the training program is highly effective across all four evaluation levels. At the reaction level, teachers showed positive responses with mean values ranging from 3.65 to 4.20, with the highest rating given to digital content quality (4.20). At the learning level, there was a significant improvement in competency with a mean difference of 20.00 points between Pre-Test and Post-Test results ($p < 0.001$) and an effect size of Cohen's $d = -5.449$, indicating an extraordinarily significant impact. At the behavior level, positive changes in teaching behavior were observed with mean values of 3.85-4.00 across four competency dimensions, demonstrating successful transfer of learning to classroom practice. At the results level, the program generated transformative impacts through the implementation of innovative methodologies including flipped classroom, interactive multimedia, gamification learning, and real-time digital assessment, resulting in improved student learning outcomes, enhanced learning resource efficiency, and optimized curriculum achievement. The findings indicate that the Platform Merdeka Mengajar self-directed training program has successfully created a more dynamic, innovative, and sustainable learning ecosystem at SMK Negeri 2 Sungai Penuh, with measurable positive impacts across all Kirkpatrick evaluation levels.

**Acknowledgments**

The authors would like to express their sincere gratitude to SMK Negeri 2 Sungai Penuh for providing access to conduct this research and for the cooperation of all teaching staff who participated in this study. Special appreciation is extended to the school administration and the teachers who generously shared their time and experiences during the data collection process. The authors also acknowledge the support of Universitas Negeri Padang for providing the necessary resources and facilities to complete this research. We are grateful to the Ministry of Education, Culture, Research, and Technology of the Republic of Indonesia for developing the Platform Merdeka Mengajar (PMM) that served as the subject of this evaluation study. Additionally, the authors thank all colleagues who provided valuable feedback and suggestions during the research process, contributing to the improvement of this study's quality and rigor.

**Author Contributions**

Conceptualization, S.S. and F.R.; methodology, S.S.; software, M.G.; validation, S.S., F.R. and S.H.; formal analysis, S.S. and M.G.; investigation, S.S. and L.S.; resources, F.R.; data curation, S.S. and M.G.; writing—original draft preparation, S.S.; writing—review and editing, F.R. and S.H.; visualization, M.G.; supervision, F.R. and S.H.; project administration, S.S.; funding acquisition, L.S. All authors have read and agreed to the published version of the manuscript.

**Funding**

This research received no external funding.

**Conflicts of Interest**

The authors declare no conflicts of interest. The research was conducted independently without any financial or personal relationships that could inappropriately influence the representation or interpretation of the reported research results. The study's design, data collection, analysis, interpretation, and the decision to publish the results were made solely by the authors without external influence.

## References


Alsalamah, A., & Callinan, C. (2021). Adaptation of Kirkpatrick's Four-Level Model of Training Criteria to Evaluate Training Programmes for Head Teachers. *Education Sciences*, *11*(3), 116. https://doi.org/10.3390/educsci11030116

Du, J. (2021). Research and Practice of Training Effect Evaluation System Based on Kirkpatrick Model-Taking Teacher Training System of Zhejiang Open University as an Example. *Chinese Studies*, *10*(2), 123-146. https://doi.org/10.4236/chnstd.2021.102009

Farida, K., & Nurhadi, D. (2023). Kirkpatrick's Evaluation Model for The Implementation of Industrial Standard GTAW Welding in Vocational Teacher Training Center. *Jurnal Inovasi Teknologi dan Edukasi Teknik*, *3*(4), 182-190. https://doi.org/10.17977/um068v3i42023p182-190

Hidayat, R., Alam, B. P., Lutvaidah, U., & Santosa, P. P. P. (2023). Implementasi Evaluasi Model Kirkpatrick Pada Program Pelatihan Penggunaan Eyd V Smk Broadcasting Mahardika. *Abdi Jurnal Publikasi*, *1*(5), 436-441.

Imansari, N., Kholifah, U., & Sasono, A. M. (2023). Evaluation of Programmable Logic Controller Training Implementation Using Kirkpatrick (4 Levels). *SAR Journal*, *6*(2). https://www.ceeol.com/search/article-detail?id=1131782

Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi. (2022, Februari). Kemendikbudristek Ciptakan Platform Merdeka Mengajar untuk Mengajar, Belajar, dan Berkarya. https://www.kemdikbud.go.id/main/blog/2022/02.

Kirkpatrick, J. D., & Kirkpatrick, W. K. (2016). *Kirkpatrick's Four Levels of Training Evaluation*. Association for Talent Development.

Lapisa, R., Lasyatta Syaifullah, Hasan Maksum, Dedy Irfan, Rifdarmon, & Iffarial Nanda. (2025). Digital Technology Innovation in TVET: Rifdarmon-Based E-Learning Model Enhancing Learning Outcomes and 4C Skills. *Salud, Ciencia Y Tecnología*, *5*, 1491. https://doi.org/10.56294/saludcyt20251491

Mahmoodi, M., Rashtchi, M., & Abbasian, G. R. (2019). Evaluation of In-Service Teacher Training Program in Iran: Focus on The Kirkpatrick Model. *Education and Self Development*, *14*(4), 20-38. https://doi.org/10.26907/esd14.4.03

Makinde, W. A., & Bamiro, T. O. (2024). Evaluation of Technical and Vocational Education and Training (TVET) Programmes in Nigeria: The Kirkpatrick Model. *Journal of Pedagogy and Education Science*, *3*(02), 126-139. https://doi.org/10.56741/jpes.v3i02.522

McLaughlin, J. A., & Jordan, G. B. (2015). Using Logic Models. In K. E. Newcomer, H. P. Hatry, & J. S. Wholey (Eds.), *Handbook of Practical Program Evaluation* (4th ed., pp. 54-87). Jossey-Bass.

Mulyasa, E. (2021). *Standar Kompetensi dan Sertifikasi Guru*. Remaja Rosdakarya.

Ramadhan, M. A., Sumarsono, R. A., Achmad, A. A., & Cisse, A. (2022). Implementation of Kirkpatrick Evaluation Model in Building Information Modeling (BIM) Training Program. *Jurnal Pendidikan Teknologi dan Kejuruan*, *28*(2), 130-138. https://doi.org/10.21831/jptk.v28i2.43555

Republik Indonesia. (2003). *Undang-Undang Nomor 20 Tahun 2003 tentang Sistem Pendidikan Nasional*.

Republik Indonesia. (2005). *Undang-Undang Republik Indonesia Nomor 14 Tahun 2005 tentang Guru dan Dosen*. Sekretariat Negara.

Romdon, Z., & Arief, Z. A. (2024). Implementasi Model Kirkpatrick dalam Evaluasi Pelaksanaan Program P5 Kewirausahaan Di SMK PGRI 3 Kota Bogor. *Prosiding Teknologi Pendidikan*, *4*(1), 100-105.

Sagala, S. (2019). *Kemampuan Profesional Guru dan Tenaga Kependidikan*. Alfabeta.

Suprihatiningrum, J. (2022). *Guru Profesional: Pedoman Kinerja, Kualifikasi & Kompetensi*. Pustaka Pelajar.

Uno, H. B. (2020). *Profesi Kependidikan: Problema, Solusi, dan Reformasi Pendidikan*. Bumi Aksara.