THE EVALUATION OF THE EFFECTIVENESS OF CAPACITY BUILDING IN MI HIDAYATUN NAJAH USING KIRKPATRICK'S MODEL

Dian Rustyawati [*1], **M. Fauzi** [2]

[1, 2,] Institut Agama Islam Nahdlatul Ulama Tuban, Indonesia

[1] ✉ awardeean@gmail.com, [2] ✉ m.fauzi@gmail.com,

Abstract	History Articles
This study focuses on evaluating the effectiveness of teacher and educational personnel training at MI Hidayatun Najah Tuban using the Kirkpatrick evaluation model, which consists of four levels: reaction, learning, behaviour, and results. The main objective of the research is to examine the extent to which the training programme improves participants' perceptions, learning outcomes, workplace behaviour, and organisational performance. The study employed a quantitative research design grounded in a positivist approach. The research population comprised all teachers and educational staff who participated in the training programme, totalling 82 individuals, and a saturated sampling technique was applied. Data were collected through questionnaires, observation, and pre-test and post-test instruments, and analysed using Likert scale analysis, descriptive statistics, and paired sample t-tests with the support of Microsoft Excel and SPSS 25.0. The findings indicate that, at the reaction level, participants showed very positive responses to the training materials, delivery methods, and relevance to their professional needs. At the learning level, statistical analysis revealed a significant increase in participants' knowledge and understanding, as evidenced by the substantial difference between pre-test and post-test scores. Behavioural evaluation results demonstrate that participants were able to transfer learning into practice, reflected in increased enthusiasm, collaboration, and improved job performance. At the results level, the training programme produced significant organisational impacts, particularly in enhancing the quality of educational services provided by the institution. The implications of this study suggest that the Kirkpatrick model is an effective and relevant framework for evaluating training in Islamic educational institutions. The results underscore the importance of systematic training and continuous evaluation to strengthen human resource development and institutional performance.	
Keywords: *Training Effectiveness, Kirkpatrick Model, Teacher Training, Islamic Educational Institutions*	
How to Cite: Rustyawati, D., & Fauzi, M. (2025). The Evaluation of The Effectiveness of Capacity Building in MI Hidayatun Najah Using Kirkpatrick's Model. *Transformation of Islamic Management and Education*, *2*(2), 75–83. Doi: https://doi.org/10.65663/timejournal.v2i2.102	

INTRODUCTION

Public and private sector organisations agree that development and training are essential to business growth (Elahi, & Bilal, 2025). Training and development offer various methods to provide new and existing teachers and educational personnel with the skills they need to perform their jobs. Training facilitates organisational expansion, improves profitability, and develops employees' capabilities (Sung, & Choi, 2023).

Organisations need to adopt strong training and improve practical training. The organisation needs to identify an effective strategy and guarantee the training and development event's success (Hayati, & Yulianto, 2021). The effectiveness of training and development was thoroughly assessed across the four levels of Kirkpatrick's Model: reaction, learning, behaviour, and results (Khosyiin, 2022).

Many organisations are unsure about the effectiveness of their training programs and development. Training represents a company's initiative to assist employees in acquiring and enhancing job-related skills (Kartika, et al., 2021). Training and development of employees is crucial for organisations to enhance service quality, boost productivity, lower labour costs, and efficiently handle workforce diversity (Gustiana, et al., 2022).

The effectiveness of a training programme involves both the acquisition of training and the transfer of knowledge (Hayati, & Yulianto, 2021). The effectiveness of training is primarily reflected in the transfer of knowledge, as demonstrated by behavioural changes

in the workplace. Training remains a crucial approach for improving individual productivity and job performance (Van Assen, 2021).

Training holds value for organisations unless it results in enhanced performance (Garavan, et al., 2021). Training effectiveness evaluates how well training achieves its objectives, such as improving job performance (Bhatt, & Muduli, 2023). Developing and implementing an effective training programme requires an evaluation of the training process (Castro, & Tumibay, 2021). Evaluating the feasibility of preparation can identify areas for training enhancement and offer insights for enhancement (Suryani, & Rindaningsih, 2023).

Most research on training effectiveness has concentrated on identifying the connections between training systems, practices, or components both individual and organizational and their overall effectiveness. This focus includes objectives, content, organizational components, applications, training duration, employee engagement, delivery methods, efficiency, innovation, and overall effectiveness results (Azevedo, et al., 2021).

Assessment is valuable for evaluating the worth, significance, or importance of an action or process implications. The Kirkpatrick Model outlines four levels of assessment: reaction, learning, behaviours, and results (Effendi, et al., 2022).

Evaluation is a crucial phase of curriculum and programme development in education. There is growing pressure to evaluate curriculum and educational programmes for various purposes, usually to assess achievement goals (Bahrissalim & Fauzan, 2018). Consequently, education integrates evaluation models from other sectors, such as business, to assess the achievement of its educational objectives. Nonetheless, the appropriateness of these evaluation models is context-dependent.

The Kirkpatrick model includes four key levels: reaction, learning, behaviour, and impact. These levels were specifically designed to evaluate workplace environments training (Khosyiin, 2022). It is recommended that all programs experience appraisal at dynamic levels, depending on available resources. Each level focuses on different aspects, as illustrated by the Kirkpatrick Model. The response level measures how satisfied participants are and their feelings about the training program. By analysing engagement, contributions, and reactions, evaluators gain valuable insights into participants' perceptions of the training program.

The learning level evaluates the range of skills, knowledge, and values participants acquire from the program. It assesses their beliefs about their ability to bring about desired changes, their confidence in executing them, and their motivation levels. In contrast, the behaviour level focuses on changes in participants' actions (Tamsuri, 2022).

This level should be measured over several weeks or months after participants complete the training program. The impact level assesses the institutional outcomes that demonstrate a positive return on investment directly associated with the training programme. Considering these institutional outcomes, developing a method to evaluate these long-term results presents a significant challenge (Gustiana, et al., 2022). The Kirkpatrick model is well-recognised in many evaluation studies for its benefits (Nawaz, et al., 2022).

MI Hidayatun Najah is one of the leading Islamic educational institutions in the Tuban district. To continually enhance its quality and services, MI Hidayatun Najah regularly conducts training and mentoring for its teachers and staff. This initiative aims to increase human resource capacity and create a supportive work environment that facilitates the development of MI Hidayatun Najah.

Several studies have been conducted at Hidayatun Najah, including research on the development of learning media (I'zaati, et al., 2022). Student Daily Report evaluation (Anisah, & Ani, 2022), Online learning models for Fiqh lessons during the COVID-19 pandemic era at MIN 1 and MIS Hidayatun Najah Tuban (Jaohari et al., 2022), Implementation of the Problem-Based Learning Model to Improve Students' Mathematics Learning Outcomes at MI Hidayatun Najah (Hikmah, et al., 2024), development of teaching materials for the subject of Al Qur'an hadith integrated with YouTube for 5th grade student MI Hidayatun Najah Tuban (Choiriyah, 2022) and the role and function of the foundation in managing education at MI Hidayatun Najah Tuban (Fauzi, 2022). This research distinguishes itself from previous studies by assessing training through the Kirkpatrick model.

RESEARCH METHODS

This study uses quantitative research methods, specifically those based on positivism, to investigate specific populations or samples. It involves data collection using research instruments and quantitative or statistical data analysis, with the objective of testing predetermined hypotheses (Azhari, et al., 2023). This study's participants were teachers and education personnel involved in the training. The population also includes 82 members. The sampling technique involves saturated samples, incorporating all population members into the research. This study employs quantitative data, sourced from both primary and secondary sources. The data collection methods used include observation and questionnaires (Asrulla, et al., 2023).

Data analysis methods focus on Likert Scale measurements using Microsoft Excel and SPSS 25.0. This involves applying the mean difference test method for two paired samples (paired sample t-test) using SPSS 25.0 software. The t-test model is utilised to analyse research models in pre-test and post-test scenarios. T-tests assess the impact of specific treatments on the same sample during two different observation periods.

FINDINGS AND DISCUSSION
Descriptive Analysis of Reaction Evaluation Variables on Training Effectiveness

In assessing the first level of the Kirkpatrick method, we observe participants' reactions following their training. The table below presents the frequency distribution of a study based on five questions, compiled from 82 participants who are engaged in this training using Microsoft Excel.

Table 1. Descriptive Analysis of Reaction Evaluation Variables on Training Effectiveness

NO.	Statement	Answers					N	Total Score	Average	TCR (%)	Category
		SD	D	N	A	SA					
		1	2	3	4	5					
1	Training materials are presented attractively	0	1	0	14	67	82	393	4,79	96	Very good
2	Training materials are beneficial for work.	0	0	2	21	59	82	385	4,70	94	Very good
3	Training materials are easy to understand.	0	0	2	30	50	82	376	4,59	92	Very good
4	Training materials are developed based on requirements	0	0	3	21	58	82	383	4,67	93	Very good
5	The training materials meet the necessary expectations and requirements.	0	0	2	24	56	82	382	4,66	93	Very good

Source: Processed from questionnaires using Excel 2021

Observations indicated that participants were actively engaged in all training sessions. Data analysis indicates that participants are generally enthusiastic about the capacity-building training at MI Hidayatun Najah. This sentiment is also reflected in a study that assesses substantial technical training concerning Learning Planning Materials in the Riau Islands Province (Dewi, et al., 2024).

Evaluating Learning Outcomes on Training Effectiveness

The second level of the Kirkpatrick method evaluates participants' learning following training (Susanty, 2022). The table below shows the results from a comparison test of pre-test and post-test scores for 82 teachers and education personnel at MI Hidayatun Najah, utilising SPSS 25.00 software.

Table 2. Paired Samples Statistic

		Mean	N	Std. Deviation	Std. Error Mean
air 1	PreTest	73.9024	82	6.62135	.73121
	Post Test	90.6707	82	5.86695	.64790

Based on the table above, the learning scores of the trainees have increased from

an initial average of 73.902 to 90.671. This indicates that there was an improvement in learning before the participants attended the training. The standard deviation value for the pre-test was 6.621, while for the post-test, it was 5.887. Finally, the standard error mean for the pre-test is 0.731, and for the post-test, it is 0.648. The average learning outcomes in the pre-test (73.902) are lower than in the post-test (90.671), indicating that the Learning Evaluation on Training Effectiveness is significant. Therefore, Training Effectiveness is considered significant.

Table 3. Paired Samples Correlations

		N	Correlation	Sig.
air 1	Pre-Test & Post-Test	82	.107	.041

Based on the table above, since the significance value (0.000) is less than α (0,05), it can be concluded that the correlation is significant. The output above shows the correlation test results, which examine the relationship between the pre-test and post-test variables. The table above indicates the correlation coefficient of 0.107, with a significance value (Sig.) 0.000. Because the Sig. Value 0.000 < probability 0.05, it can be stated that a relationship exists between the pre-test variable and the post-test variable.

Table 4. Paired Samples Test

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
Pair 1 PreTest - Post Test	-16.76829	8.36557	.92382	-18.60641	-14.93018	-18.151	81	.000

Based on the table above, the mean column shows an average difference of -16.768 before and after training. This value reflects the difference between the average outcomes of the pre-test and those of the post-test. The Standard Deviation column shows the standard deviation of the average differences between pre-test and post-test scores of trainees, which is 8.366. The Standard Error Mean column is the variability index (diversity) 0.924. The t column shows a negative t count, indicating that the average pre-test learning outcomes are lower than the average post-test outcomes. As a result, the negative t value is modified to a positive value of 18.151.

To find the value of the t-table, the t-table is looked for based on the value of df (degrees of freedom) and the significance value ($α/2$). The table above shows that the df value is 81, and the value of 0.05/2 equals 0.025. Thus, the value of the t-table is 1.98969. Given that the calculated t value is 18.151 > t-table 1.98969, it can be concluded, based on the decision-making criteria outlined above, that H_0 is rejected and H_a is accepted. Therefore, there is an average difference in the learning outcomes of training in the pre-test and post-test among the teachers and education personnel of MI Hidayatun Najah Tuban.

Descriptive Analysis of Behavioural Evaluation Variables on Training Effectiveness

In testing the third level of the Kirkpatrick method, namely, observing participants' behaviour after attending training (Azizah, & Fatonah, 2023), the following table presents a frequency distribution table of research consisting of 4 questions and the frequency distribution of data processing from 82 respondents engaged in this training using Microsoft Excel.

Table 5. Descriptive Analysis of Behavioural Evaluation Variables on Training Effectiveness

NO.	Statement	Answers					N	Total Score	Average	CR (%)	Category
		DD	D	N	A	SA					
		1	2	3	4	5					
1	I understand the material and I am more enthusiastic about my work.	0	0	3	34	45	82	361	4,40	88	Good
2	I can collaborate with fellow teachers and education personnel to develop the institution.	0	0	6	34	42	82	364	4,44	89	Good
3	I can apply the things learned during the training in my work environment.	0	2	0	27	53	82	377	4,60	92	Very Good

4	I feel there is an increase in my ability to do my job and provide the best educational services in the institution.	0	2	0	24	56	82	380	4,63	93	Very Good

According to the table above, two statements yield good results, and two statements yield excellent results. Therefore, it is evident that the evaluation results of participants' behaviour after the training have improved. This capacity-building training for teachers and education personnel has successfully introduced positive vibes at MI Hidayatun Najah. This aligns with research on the Kirkpatrick evaluation model, which emphasises training future madrasa leaders at the technical personnel training for Education and Religious Education centre, where the analysis results indicate excellent qualifications. After the training, participants became more polite and friendly (Engriyani, & Rugaiyah, 2022).

Evaluation Assessment of Training Effectiveness

In this test, the fourth in the Kirkpatrick method is utilised to assess the participants' results after attending the training (Sukeriyadi, 2024). The following table presents the results of the comparative analysis between the pre-test and post-test training, as observed from the impact on the organisation of the participants' work to enhance the performance of teachers and education personnel in providing educational services using SPSS 25.00 software.

Table 6. Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
air 1	Pre test	78.2927	82	5.56774	.61485
	Post Test	91.5854	82	6.61452	.73045

Based on the table above, the learning scores of the trainees have risen from an initial average of 78.293 to 91.585. This indicates that there was an improvement in learning before the participants attended the training. The standard deviation value for the pre-test was 5.568, whereas for the post-test, it was 6.61. Finally, the standard error mean for the pre-test is 0.615, and for the post-test, it is 0.730. The average learning outcomes in the pre-test (78.293) are lower than in the post-test (91.585), suggesting that the evaluation assessment of training effectiveness is significant.

Table 1 Paired Samples Correlations

		N	Correlation	Sig.
air 1	Pretest & Post Test	82	.060	.594

Based on the table above, as the significance value (0.000) is less than α (0,05), it can be concluded that the correlation is significant. The output above presents the correlation test results, which examine the relationship between the pre-test and post-test variables. The table above indicates that the correlation coefficient is 0.060, with a significance value (Sig.) of 0.000. Since the Sig. Value of 0.000 is less than the probability of 0.05, it can be stated that a relationship exists between the pre-test variable and the post-test variable.

Table 2 Paired Sample Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	Pretest-Post Test	-13.29268	8.89651	.98245	-15.24746	-11.33791	-13.530	81	.000

Based on the table above, the mean column shows an average difference of -13.293 before and after training. This value reflects the difference between the average outcomes of the pre-test and those of the post-test. The Standard Deviation column shows the standard deviation of the average differences between pre-test and post-test scores of trainees, which is 8.897. The Standard Error Mean column is the variability index (diversity) 0.982. The t column shows a negative t count, indicating that the average pre-test learning outcomes are lower than the average post-test outcomes. As a result, the negative t value is modified to a positive value of 13.530.

To determine the value of the t-table, one refers to it based on the degrees of freedom (df) and the significance value (α/2). The table above indicates that the df value is 81, and the value of 0.05/2 equals 0.025. Therefore, the value of the t-table is 1.98969.

Given that the calculated t value is 13.530 > t-table 1.98969, it can be concluded, by the evaluation assessment criteria outlined above, that H_o is rejected and H_a is accepted. There is a significant average difference in the evaluation assessment of training in the pre-test and post-test among the teachers and education personnel of MI Hidayatun Najah Tuban. This is supported by the results of research conducted by Sudipta Paul, which used cross-sectional primary data generated through a household survey and personal interviews with one thousand randomly selected trainees to test and validate our proposed protocol. By applying it to our evaluation target, we find that the degree of effectiveness of the training programs varies (Durlak, et al., 2022). The Kirkpatrick model remains useful, relevant, and applicable across various contexts. It can be tailored to different training environments and excels at evaluating training effectiveness (El-Sabagh, 2021).

The findings of this study demonstrate that the training programme implemented at MI Hidayatun Najah Tuban is effective when evaluated using the four levels of the Kirkpatrick model. At the reaction level, participants' responses indicate a very positive perception of the training materials, delivery, and relevance to their professional needs. High average scores and very good categories across all reaction indicators suggest that the training was well-designed and aligned with participants' expectations. This positive reaction is important, as favourable perceptions of training content and facilitation are often a prerequisite for deeper engagement and successful learning outcomes. Consistent with previous studies, positive reactions reflect participants' readiness to absorb new knowledge and skills, which strengthens the overall effectiveness of the training process.

At the learning level, the paired sample t-test results reveal a statistically significant increase in participants' knowledge and understanding after attending the training. The substantial difference between pre-test and post-test scores confirms that the training programme successfully enhanced cognitive learning outcomes among teachers and educational personnel. This improvement indicates that the training content, methods, and instructional strategies were effective in facilitating knowledge acquisition. The significant correlation between pre-test and post-test scores further supports the conclusion that the observed changes are attributable to the training intervention rather than random variation. These results reinforce the argument that structured and systematic training can meaningfully improve teachers' professional competencies.

The behavioural evaluation shows that learning outcomes were transferred into workplace practices. Participants reported increased enthusiasm for work, improved collaboration with colleagues, and greater ability to apply training materials in their daily professional activities. The predominance of good and very good categories across behavioural indicators suggests that the training not only increased knowledge but also influenced attitudes and work-related behaviours. This finding is particularly important, as behaviour change is often considered a critical indicator of training success, reflecting the extent to which learning is applied in real work contexts. In the context of MI Hidayatun Najah, these behavioural changes contribute to a more collaborative, motivated, and service-oriented institutional culture.

At the results or impact level, the study confirms that the training programme produced significant organisational benefits. The statistically significant difference between pre-test and post-test scores related to performance outcomes indicates that training contributed to improved institutional performance in delivering educational services. Although measuring long-term organisational impact remains a challenge, the findings suggest that training has positively influenced teachers' and staff members' effectiveness, which in turn supports institutional development. These results align with previous research highlighting the continued relevance and adaptability of the Kirkpatrick model in assessing training outcomes across different organisational and educational settings.

The implications of this study are both theoretical and practical. Theoretically, the findings strengthen the empirical support for the use of the Kirkpatrick model in Islamic educational institutions, demonstrating its applicability beyond corporate or industrial training contexts. This study contributes to the literature by providing evidence that the model can comprehensively capture training effectiveness in madrasah-based educational environments. Practically, the results imply that school leaders and policymakers should continue investing in systematic training and development programmes, as these initiatives have proven benefits at individual, behavioural, and organisational levels. Furthermore, the positive outcomes suggest that training programmes should be designed with clear objectives, relevant content, and appropriate

evaluation mechanisms to ensure sustainable impact. By integrating regular evaluation using established models such as Kirkpatrick, educational institutions can continuously improve the quality of human resources and enhance overall institutional performance.

CONCLUSION

The findings demonstrate that the training and mentoring programme at MI Hidayatun Najah Tuban is effective across all four levels of the Kirkpatrick evaluation model. At the reaction level, participants expressed very positive responses to the training materials, delivery, and relevance to their professional needs, indicating high satisfaction and engagement. At the learning level, the paired sample t-test results show a significant increase in participants' knowledge and understanding, as reflected in the substantial difference between pre-test and post-test scores. These results imply that the training programme successfully enhanced teachers' and educational personnel's competencies, confirming that well-designed capacity-building initiatives can meaningfully strengthen human resources in Islamic educational institutions.

At the behaviour and results levels, the study found positive changes in workplace practices and institutional performance. Participants reported increased enthusiasm, collaboration, and ability to apply newly acquired knowledge in their daily tasks. Moreover, the significant differences between pre-test and post-test results at the impact level indicate that the training contributed to improved organisational outcomes and service quality. These findings imply that systematic training evaluation using the Kirkpatrick model can provide evidence-based insights for school management, enabling leaders to align professional development programmes with institutional goals and ensure a positive return on investment in human resource development.

Despite these contributions, this study has several limitations. First, it relies primarily on quantitative self-reported data, which may not fully capture deeper behavioural changes or long-term institutional impacts. Second, the research focuses on a single institution, limiting the generalisability of the findings to other educational contexts. Additionally, the study evaluates training outcomes within a relatively short time frame, leaving gaps in understanding the sustainability of training effects. Future research is therefore recommended to employ mixed-methods approaches, incorporate longitudinal designs, and involve multiple institutions. Such efforts would strengthen the validity of findings, address existing gaps, and provide a more comprehensive evaluation framework for training effectiveness in Islamic and general educational institutions.

ACKNOWLEDGMENT

REFERENCES

Anisah, Z., & Ani, R. A. (2022). Daily Report: Sebagai Evaluasi Kedisiplinan Siswa Kelas III MI Hidayatun Najah Tuban. *BASICA*, *2*(1), 29–41. https://doi.org/10.37680/basica.v2i1.1520

Asrulla, R., Jailani, M. S., & Jeka, F. (2023). Populasi dan sampling (kuantitatif), serta pemilihan informan kunci (kualitatif) dalam pendekatan praktis. *Jurnal Pendidikan Tambusai*, *7*(3), 26320–26332.

Azevedo, M. C. D., Schlosser, F., & McPhee, D. (2021). Building organizational innovation through HRM, employee voice and engagement. *Personnel Review*, *50*(2), 751-769. https://doi.org/10.1108/PR-12-2019-0687

Azhari, M. T., Al Fajri Bahri, M. P., Asrul, M. S., & Rafida, T. (2023). *Metode penelitian kuantitatif*. PT. Sonpedia Publishing Indonesia.

Azizah, D. N., & Fatonah, U. (2023). KONSEP MODEL KIRKPATRICK DALAM EVALUASI PROGRAM PELATIHAN. *PROSIDING TEKNOLOGI PENDIDIKAN*, *3*(1), 69–74.

Bahrissalim, B., & Fauzan, F. (2018). Evaluasi Kurikulum Pelatihan Dalam Meningkatkan Kompetensi Pedagogik Guru Pai Di Balai Diklat Keagamaan Jakarta. *Edukasia: Jurnal Penelitian Pendidikan Islam*, *13*(1), 25–52. https://doi.org/10.21043/edukasia.v13i1.2779

Bariqi, M. D. (2018). Pelatihan dan pengembangan sumber daya manusia. *Jurnal studi*

manajemen dan bisnis, *5*(2), 64–69.
https://doi.org/10.21107/jsmb.v5i2.6654

Bhatt, P., & Muduli, A. (2023). Artificial intelligence in learning and development: a systematic literature review. *European Journal of Training and Development*, *47*(7/8), 677-694.
https://doi.org/10.1108/EJTD-09-2021-0143

Castro, M. D. B., & Tumibay, G. M. (2021). A literature review: efficacy of online learning courses for higher education institution using meta-analysis. *Education and Information Technologies*, *26*(2), 1367-1385.
https://doi.org/10.1007/s10639-019-10027-z

Choiriyah, S. U. (2022). Pengembangan Bahan Ajar Mata Pelajaran Al-Quran Hadis yang Terintegrasi Youtube Bagi Peserta Didik Kelas 5 Madrasah Ibtidaiyah Hidayatun Najah Tubah. *EDU-RELIGIA: Jurnal Keagamaan dan Pembelajarannya*, *5*(1), 43-59.
https://doi.org/10.52166/edu-religia.v5i1.2976

Dewi, N. P., Indrayani, I., Wibisono, C., & Rumengan, A. E. (2024). Government Employee Assessment to Support Public Performance Organization Learned from Riau Island Province. *Cogent Business & Management*, *11*(1), 1-18.
https://doi.org/10.1080/23311975.2024.2368322

Durlak, J. A., Mahoney, J. L., & Boyle, A. E. (2022). What we know, and what we need to find out about universal, school-based social and emotional learning programs for children and adolescents: A review of meta-analyses and directions for future research. *Psychological Bulletin*, *148*(11-12), 765.
https://doi.org/10.1037/bul0000383

Effendi, M., Zainuddin, Z., & Ahmad, M. S. (2022). Implementasi evaluasi model kirkpatrick terhadap kualitas sistem pembelajaran. *Jurnal EDUCATIO: Jurnal Pendidikan Indonesia*, *8*(1), 1–8.
https://doi.org/10.29210/1202221160

Elahi, F., & Bilal, A. R. (2025). Elevating excellence: a standardized process of training and development for private sector educational SMEs. *Business Process Management Journal*, *31*(3), 1123-1148.
https://doi.org/10.1108/BPMJ-10-2023-0806

El-Sabagh, H. A. (2021). Adaptive e-learning environment based on learning styles and its impact on development students' engagement. *International Journal of Educational Technology in Higher Education*, *18*(1), 53.
https://doi.org/10.1186/s41239-021-00289-4

Engriyani, E., & Rugaiyah, R. (2022). Evaluasi program pendidikan dan pelatihan menggunakan Model Kirkpatrick pada Institusi Pemerintah di Indonesia. *Jurnal Ilmiah Wahana Pendidikan*, *8*(23), 683–692.

Fauzi, M. (2022). PERAN DAN FUNGSI YAYASAN DALAM PENGELOLAAN PENDIDIKAN MADRASAH IBTIDAIYAH HIDAYATUN NAJAH TUBAN. *Eduthink: Jurnal Pemikiran Pendidikan Islam*, *3*(02).

Garavan, T., McCarthy, A., Lai, Y., Murphy, K., Sheehan, M., & Carbery, R. (2021). Training and organisational performance: A meta-analysis of temporal, institutional and organisational context moderators. *Human Resource Management Journal*, *31*(1), 93-119.
https://doi.org/10.1111/1748-8583.12284

Gustiana, R., Hidayat, T., & Fauzi, A. (2022). Pelatihan dan pengembangan sumber daya manusia (suatu kajian literatur review ilmu manajemen sumber daya manusia). *Jurnal Ekonomi Manajemen Sistem Informasi*, *3*(6), 657–666.

Hayati, N., & Yulianto, E. (2021). Efektivitas pelatihan dalam meningkatkan kompetensi sumber daya manusia. *Journal Civics and Social Studies*, *5*(1), 98–115.
https://doi.org/10.31980/civicos.v5i1.958

Hikmah, A. F., Faiziyah, N., & Ishartono, N. (2024, January). Implementation of etnomathematics-oriented problem-based learning to improve student activities and learning outcomes. In *AIP Conference Proceedings* (Vol. 2926, No. 1, p. 020020). AIP Publishing LLC.
https://doi.org/10.1063/5.0183046

I'zaati, L., Pradana, A. A., & Ma'rifati, R. K. D. N. (2022). Pengembangan Media Komik Dalam Pembelajaran Bahasa Indonesia Tema Cita-Citaku Di Mi Hidayatun Najah Tuban. *PREMIERE: Journal of Islamic Elementary Education*, *4*(1), 29–40.
https://doi.org/10.51675/jp.v4i1.274

Jaohari Farchan, T. (2022). *Model Pembelajaran Daring Pendidikan Agama Islam Melalui Materi Wudhu dalam Masa Pandemi Covid-19 di Mts Tholabuddin Masin Warungasem Kabupaten Batang* (Doctoral dissertation, IAIN Pekalongan).

Kartika, M., Khoiri, N., Sibuea, N. A., & Rozi, F. (2021). Learning by doing, training and life skills. *MUDABBIR Journal Research and Education Studies*, *1*(2), 91–103. https://doi.org/10.56832/mudabbir.v1i2.80

Khosyiin, M. I. (2022). Evaluasi program pelatihan model kirkpatrick. *CERMIN: Jurnal Manajemen Dan Pendidikan Berbasis Islam Nusantara*, *1*(2), 42–46.

Nawaz, F., Ahmad, W., & Khushnood, M. (2022). Kirkpatrick model and training effectiveness: a meta-analysis 1982 to 2021. *Business & Economic Review*, *14*(2), 35–56. https://doi.org/10.22547/BER/14.2.2

Noe, R. A. (2020). *Employee training and development*. McGraw-Hill.

Sukeriyadi, M. (2024). Evaluasi Program Pendidikan: Training Program Evaluation Model Krikpatrick. *NABAWI: Jurnal Penelitian Pendidikan Islam*, *2*(1), 101–117.

Sung, S. Y., & Choi, J. N. (2023). What drives firms to invest in training and developing employees? Time-dependent effects of firm internal and external contingencies. *The InTernaTIonal Journal of human resource managemenT*, *34*(2), 223–252. https://doi.org/10.1080/09585192.2021.1965007

Suryani, S., & Rindaningsih, I. (2023). Pelatihan dan pengembangan sumber daya manusia. *PERISAI: Jurnal Pendidikan dan Riset Ilmu Sains*, *2*(3), 363–370.

Susanty, Y. (2022). Evaluasi program pengembangan kompetensi berdasarkan model evaluasi Kirkpatrick level 1 dan level 2. *Jurnal Administrasi Publik*, *18*(2), 172–191. https://doi.org/10.52316/jap.v18i2.111

Tamsuri, A. (2022). Literatur review penggunaan metode kirkpatrick untuk evaluasi pelatihan di Indonesia. *Jurnal Inovasi Penelitian*, *2*(8), 2723–2734.

Van Assen, M. F. (2021). Training, employee involvement and continuous improvement–the moderating effect of a common improvement method. *Production planning & control*, *32*(2), 132-144. https://doi.org/10.1080/09537287.2020.1716405