

The Developing of Kirkpatrick's Evaluation Model on the Evaluation of Early Childhood Head Training in Early Children Education Development Center for Early Children and Community Education North Sumatera

Kristianus Mote
{mote.kris@yahoo.com}

Doctoral Student State University of Medan, Indonesia

Abstract. This study aims to determine the effectiveness of the training of the head of the early childhood education unit by applying the Kirkpatrick's evaluation model and to determine the effectiveness of the training of the head of the early childhood education unit by applying other evaluation models. The research method used is a mixed method. Data collection techniques are observation and tests. The research sample was 94 heads of early childhood education units. The data analysis technique used is descriptive analysis. The research model test is the N-Gain Score test. The results showed that the effectiveness of the implementation of the Kirkpatrick's evaluation model showed 0.42% in the medium category, while the Mote's evaluation model based on the national education system and based on statistical methods showed 0.70% in the high category. Therefore, it is recommended to use the Mote's evaluation model that so the high result.

Keywords: developing, kirkpatrick's evaluation model, head of early childhood

1 Introduction

In the process of managing the quality of human resources through various forms of activities including training activities, the management function is known, one of which is the evaluation function. According to Stark, J.S. & Thomas, A. [1] that evaluation is a process or activity of selecting, collecting, analyzing and presenting information that can be used as a basis for decision making and further programming. Furthermore Stufflebeam, D.L. & Shinkfield, A.J. [2] says that evaluation is a process of providing information that can be used as consideration for determining prices and services from the objectives achieved, design, implementation and impact to help make decisions, assist accountability and increase understanding of the phenomenon. According to Kaswan [3] that training is the process of increasing the knowledge and skills of employees. Furthermore Meldona & Siswanto [4] says training is a systematic process of changing the behavior of employees in a direction to increase efforts to achieve organizational goals.

According to S. Eko Putro Widoyoko [5] says that an appropriate evaluation model as needed is needed in the process of managing the quality of human resources. Kirkpatrick's

evaluation model or the four level model is recognized as having advantages because it is comprehensive, simple and can be applied in various trainings. Comprehensive in the sense that the evaluation model is able to reach all sides of the training program. Simple in the sense that the evaluation model has a logical flow that is easy to understand and clear and uncomplicated categorization. It can be applied in various trainings in the sense that the evaluation model can be used to evaluate various types of training in various situations. Kirkpatrick's evaluation model or four-level model, namely reaction evaluation, learning evaluation, behavior evaluation and outcome evaluation.

According to Kirkpatrick, Donal L. & James D. Kirkpatrick [6] says that reaction evaluation is also called level one evaluation which aims to determine the level of satisfaction of training participants with the implementation of the training, namely the effectiveness of the training process which is realized by feeling fun, interesting and motivating. There are several aspects to measure the level of satisfaction of training participants, namely the service of the organizing committee, quality of instructors, materials, methods, class atmosphere, main and supporting facilities, value and meaning and others. The tool used to measure the reactions of the trainees is in the form of a questionnaire instrument. In compiling an instrument to measure the reaction of the ideal trainee from providing the maximum amount of information and requiring the minimum amount of time. According to Center Partners [7] says that the number of items includes no more than 15-25 questions, which are designed to obtain qualitative and quantitative data.

According to Kirkpatrick, Donal L. & James D. Kirkpatrick [8] says that learning evaluation is also called level two evaluation which aims to determine learning outcomes, namely cognitive, affective and psychomotor when participants return to work. There are several aspects to measure learning outcomes, namely comparing cognitive, affective and psychomotor outcomes before and after training. The tools used to measure the learning outcomes of the trainees are in the form of test instruments, questionnaires, interviews, observations and performance appraisal rubrics. Tests were used to measure the level of improvement in knowledge, questionnaires and interviews were used to measure attitudes and observations and a performance appraisal rubric was used to measure skill levels. To determine the level of improvement in these aspects, tests were carried out before and after the training. The use of comparison groups as a reference for the effect of training on participants. The control group if practical, evaluates knowledge, skills and/or attitudes both before and after the program, paper and pencil tests to measure knowledge and attitudes and performance tests to measure skills. This comparison group is a group that did not participate in the training. The group that participated in the training and the group that did not participate in the training were compared the development of knowledge, attitudes and skills over a certain period of time. Both groups were measured and the results of the two measurements were compared to determine the effect of the training program on the participants.

According to S. Eko Putro Widoyoko [9] says that behavioral evaluation is also called level three evaluation which aims to find out changes in behavior that occur after participants return to the workplace. There are several aspects to measure behavior, namely changes in mental attitudes, improving knowledge and adding skills. The tools used to measure behavior are in the form of observations and interviews as well as document analysis. The implementation of the behavior evaluation is carried out first by giving a pause for the transition period at the latest three months after the training. Behavioral evaluation is carried out more than once in a

sufficient time span to determine whether the change in behavior is temporary or permanent. Changes in behavior after returning to work are called outcomes.

According to S. Eko Putro Widoyoko [10] says that results evaluation is also called level four evaluation which aims to determine the impact of changes in the work behavior of training participants on the level of organizational productivity that occurs because participants have attended training. There are several aspects to measure the evaluation of the results, namely the work atmosphere, the level of work participation, the quality of work and others. The tools used to measure the results are in the form of observations and interviews as well as document analysis. In general, training materials do not have a direct impact on organizational results, besides that the calculation of aspects of an organization's results is carried out in the annual report period, so this level four evaluation is difficult compared to the previous level evaluation so it requires a longer time span than behavioral evaluation. Level four evaluation is also called evaluation of training impact.

According to Brikerhoff, RD. Brethower, DM, Hluchyj. T., et al. [11] said that evaluation in training consists of seven steps as follows: 1) determine the focus to be evaluated. 2) develop evaluation design. 3) collect information. 4) analyze and interpret information. 5) make a report. 6) evaluation management. 7) evaluate for evaluation.

According to Lincoln & Arifin Zainal [12] said that evaluation in training consists of six components as follows: 1) Achievement and Accuracy of Training Objectives. In the evaluation, there must be a collection of information about the achievement and accuracy of the targets. This means whether the training has achieved the expected goals or not and whether these goals are in accordance with the training needs or not. 2) Training Materials. In the evaluation, there must be a collection of information about the material discussed during the training, including: a) Whether the material is in accordance with the objectives or not. b) Is the training material too simple, too difficult, too theoretical and so on. 3) Training Facilitator. In the evaluation, there should be a collection of information about "facilitators" who assist in the learning process. In this case, it is necessary to collect information about the facilitator's ability to facilitate training. Matters that need to be evaluated include: a) Mastery and ability to use participatory methods. b) Mastery and understanding of training materials. c) Ability to communicate and interact with participants effectively. d) Facilitation team collaboration. e) Ability to use training media effectively. 4) Training Participants. In the evaluation, there must be a collection of information about the level of participation of participants, cooperation between participants, cooperation between participants and facilitators. In addition, the criteria for participants, whether the participants involved in the training meet expectations, as set out in the terms of reference for the training and so on. 5) Effectiveness of Training Methods. In the evaluation, there should be a collection of information about the effectiveness of the method. Is the method used to encourage participant involvement, is the method used in accordance with the objectives, is the method used in accordance with the content of the training material. 6) Training Committee. In the evaluation, it is more focused on evaluating the logistical aspects. Matters that need to be evaluated include: a) Communication, which is about how notification or invitation, is one type of evaluation of the organizer, whether the invitation is clear and accompanied by the required information, equipped with a training reference frame. b) Training support facilities and infrastructure which include basic training, for plenary discussions and group discussions, consumption, accommodation, availability and readiness of materials needed by participants, facilitators, committees and so on.

Based on preliminary research shows that the training of the head of the early childhood education unit using the Kirkpatrick's evaluation model, the results of the training are less than high.

Therefore, through this training, researchers conducted research on Kirkpatrick's evaluation model or four level model, namely reaction evaluation, learning evaluation, behavior evaluation and result evaluation in order to find out and examine and provide solutions.

2 Research Methods

The research period is 36 months, starting from July 2017 to July 2020. The research site is at the Center for Early Childhood Development and Community Education in North Sumatera. This study aims: (1) to determine the effectiveness of the training of the head of the early childhood education unit by applying the Krikpatrick's evaluation model or the four level evaluation model, namely reaction evaluation, learning evaluation, behavior evaluation and result evaluation. (2) to find out the effectiveness of training for the head of the early childhood education unit by applying other evaluation models. Research method is mixed methods. Research approach is research and development. According to Nuryani Y. Rustaman, et.al. [13] says that the data collection techniques are observation and tests. Furthermore Sugiyono [14] says that the population of study were 140 head of early childhood education units. Furthermore Suharsimi Arikunto [15] says that the sample of study was 94 head of early childhood education units. Furthermore Zainuddin & Ghodang, H. [16] says that the research technique is descriptive analysis technique. Furthermore Meltzer [17] says that the test of research model is the N-Gain Score. Furthermore Ngalim Purwanto [18] says that the attendance test of the participants in study was the Science Process Test.

3 Result and Discussion

The Analysis of research data obtained 2 (two) results as follows: 1) Training Evaluation. The training evaluation used Kirkpatrick's evaluation model or four level model, namely reaction evaluation, learning evaluation, behavior evaluation and outcome evaluation. Results the data presented include the results of observation and test data. (a) Description of observation data. The description of the data is presented in Table 1.

Table 1. Observation Results

No	Indicator	Experiment Class		Control Class	
		Percentage	Description	Percentage	Description
1	Observation	72,69	Enough	70,56	Enough
2	Prediction	74,38	Enough	72,66	Enough
3	Classifica- tion	73,85	Enough	71,98	Enough
4	Communica- tion	81,67	Good	74,65	Enough

Based on Table 1 shows that there are differences in the results of observations between the experimental class and the control class in training with the indicator as follows: 1) for the

experimental class, the highest indicator, namely communication, is obtained by a percentage of 81.67% with the good category, while the lowest indicator, namely observation, is obtained, the percentage is 72.69% with the enough category. 2) for the control class, the highest indicator, namely communication, was obtained by a percentage of 74.65% with the enough category, while the lowest indicator, namely classification, obtained a percentage of 70.56% with the enough category. b) Description of test data. The description of the data is presented in Table 2.

Table 2. Test Results

N	Experiment Class				Control Class			
	Pre Test	Post Test	N-Gain	Category	Pre Test	Post Test	N-Gain	Category
Amount of Participants	47 Head				47 Head			
Average Value	72,39	84,11	0,42	Medium	72,90	82,75	0,36	Medium

Based on Table 2 shows that the average pre test value for the experimental class is 72.39% and the post test average value is 84.11% and the N-Gain is 0.42% with the medium category, while the average pre test value for the control class is 72.90% and the post test value is 82.75% and the N-Gain is 0.36% with the medium category. From the achievement of the two N-Gain in the experimental class and the control class, it is known that the N-Gain value in the experimental class is higher than the N-Gain value in the control class, meaning that learning using Kirkpatrick's evaluation model or four level model, namely reaction evaluation, evaluation learning, evaluation behavior and evaluation results for the experimental class can improve cognitive with the medium category. 2) Evaluation of the Developed Model. In training evaluation, the Mote's evaluation model or five level model is used, namely reaction evaluation, learning evaluation, behavior evaluation, result evaluation and benefit evaluation. The results of the data presented include the results of observation and test data. (a) Description of observation data. The description of the data is presented in Table 3.

Table 3. Observation Results

No	Indicator	Experiment Class		Control Class	
		Percentage	Description	Percentage	Description
1	Observation	82,97	Good	74,89	Enough
2	Prediction	89,36	Very Good	73,40	Enough
3	Classification	80,85	Good	71,98	Enough
4	Communication	82,55	Good	83,82	Good

Based on Table 3 shows that there are differences in test results between the experimental class and control class in training with the indicator as follows: 1) for the experimental class the highest indicator is obtained, namely prediction, the percentage is 89.36% with the very good category, while the lowest indicator is classification, obtained a percentage of 80.85% with the good category. 2) for the control class obtained the highest indicator, namely communication, obtained a percentage of 83.82% with the good category, while the lowest indicator, namely

classification, obtained a percentage of 71.98% with the enough category. (b) Description of test data. The description of the data is presented in Table 4.

Table 4. Test Results

N	Experiment Class			Control Class				
	Pre Test	Post Test	N-Gain	Cate Gory	Pre Test	Post Test	N-Gain	Cate gory
Amount of Participants			47 Head				47 Head	
Average Value	60,49	87,80	0,70	High	60,80	77,21	0,40	Medium

Based on Table 4 shows that the average pre test value for the experimental class is 60.49% and the post test average value is 87.80% and the N-Gain is 0.70% with the high category, while the average pre test value for the control class is 60.80% and the post test average is 77.21% and the N-Gain is 0.40% with the medium category. The pre test value of the experimental class and the control class is not much different, while the post test value for the experimental class is higher than the control class, as well as the N-Gain value for the experimental class is higher than the N-Gain value for the control class, meaning learning by using the Mote's evaluation model or the five level model, namely reaction evaluation, learning evaluation, behavior evaluation, result evaluation and benefit evaluation for the experimental class can improve cognitive with a high category.

This research is discussed as follows: 1) Training Evaluation. The results of the observation and evaluation test of training using the Kirkpatrick evaluation model or four level model, namely reaction evaluation, learning evaluation, behavior evaluation and result evaluation showed 0.42% was in the medium category. This shows that the evaluation of moderate training results has the following characteristics: (a) understanding of reaction evaluation, evaluation of learning, evaluation of behavior and evaluation of training results is in the medium category. (b) willingness to try to improve the evaluation of reactions, evaluation of learning, evaluation of behavior and evaluation of training results including the medium category. 2) Evaluation of the Developed Model. The results of observations and training evaluation tests using the Mote evaluation model or five level model, namely reaction evaluation, learning evaluation, behavior evaluation, outcome evaluation and benefit evaluation showed that 0.70% was in the high category. This shows that the evaluation of high training outcomes has the following characteristics: (a) understanding of reaction evaluation, learning evaluation, behavior evaluation, evaluation of results and evaluation of training benefits are included in the high category. (b) willingness to try to improve the evaluation of reactions, evaluation of learning, evaluation of behavior, evaluation of results and evaluation of the benefits of training included in the high category. The evaluation model is presented in Figure 1.

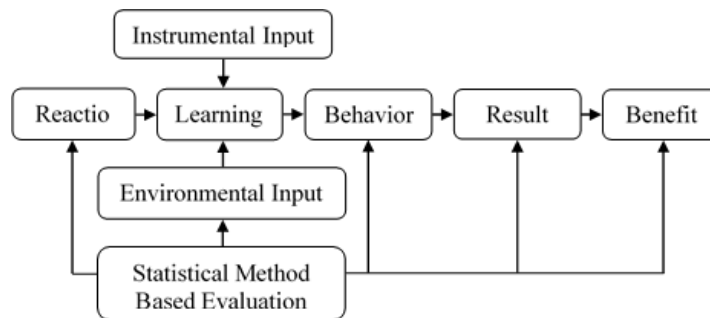


Fig.1. Mote's Evaluation Model or Five Level Model Based on the National Education System

Description:

In the training, there are five (5) stages of evaluation based on the national education system, namely: 1) reaction evaluation is input. 2) learning evaluation is process. 3) behavior evaluation is output. 4) result evaluation is outcome. 5) benefit evaluation is income. The five stages of evaluation each have aspects to be evaluated, namely 1) reaction evaluation is input. At the reaction evaluation stage, the aspects evaluated include the service of the organizing committee, the quality of the instructors, training materials, training methods, classroom atmosphere, main facilities and supporting facilities, the value and significance of the training content and others. The tool used to measure the reactions of the trainees is in the form of a questionnaire instrument. 2) learning evaluation is process. At the learning evaluation stage, the aspects that evaluated include cognitive, affective and psychomotor before and after training. The tools used to measure the learning outcomes of the trainees are in the form of test instruments, questionnaires, interviews, observations and performance appraisal rubrics. When evaluating learning, it is supported by instrumental input and environmental input. 3). behavior evaluation is output. At the behavioral evaluation stage, the aspects that are evaluated include changes in mental attitudes, improving knowledge and adding skills. The tools used to measure behavior are in the form of observations and interviews as well as document analysis. 4) result evaluation is outcome. At the result evaluation stage, the evaluated aspects include work atmosphere, work participation rate, work quality and others. The tools used to measure the results are in the form of observations and interviews as well as document analysis. 5) benefit evaluation is income. At the benefit evaluation stage, the aspects evaluated include independence, standardization, local excellence, international level orientation. The tools used to measure the results are in the form of observations, interviews, questionnaires, tests and document analysis.

The five stages of evaluation are reaction evaluation, learning evaluation, behavior evaluation, result evaluation and benefit evaluation, when conducting mandatory evaluations based on statistical methods.

4 Conclusion

This study was concluded as follows: 1) Kirkpatrick's evaluation model or four level model, namely reaction evaluation, learning evaluation, behavior evaluation and result evaluation showed that 0.42% with the medium category. 2) The evaluation model developed, namely the Mote's evaluation model or five level model, namely reaction evaluation, learning

evaluation, behavior evaluation, result evaluation and benefit evaluation showed that 0.70% with the high category.

Acknowledgement

In the process of developing this model training evaluation, various elements have been guided, so i would like to thank the following elements: Reviewers, Nasrun as Professor of the Medan State University Postgraduate Program, Benjamin Situmorang as Professor of the Postgraduate Program at the State University of Medan.

Therefore, the model is expected to be useful for all related elements in the training process in the context of developing human resources.

References

- [1] Stark, J.S. & Thomas, A. (1994). *Assesment and Program Evaluation*. Needham Heights: Simon & Schuster Custom Publishing, p. 12.
- [2] Stufflebeam, D.L. & Shinkfield, A.J. (1985). *Systematic Evaluation*. Boston: Kluwer Nijhof Publishing, p. 159.
- [3] Kaswan. (2011). *Pelatihan dan Pengembangan*. Bandung: Alfabeta. Hlm. 2. Meldona dan Siswanto. (2012). *Perencanaan Tenaga Kerja Tinjauan*
- [4] *Integratif*. Malang: UIN-Maliki Press. Hlm. 217. S. Eko Putro Widoyoko. (2017). *Evaluasi Program Pelatihan*. Yogyakarta: Pustaka Pelajar. Hlm. 169-170.
- [5] Kirkpatrick, Donal L. & James D. Kirkpatrick. (2009). *Evaluating Training Programs*. San Fransisco: Berret-Koehler Publisher Inc. p. 26. Center Partners. (2016). Implementing the Kirkpatrick Evaluation *Model Plus.*, diambil pada 2 Januari 2016, dari <http://www.coe.wayne.edu/val/pdf>.
- [6] Kirkpatrick, Donal L. & James D. Kirkpatrick. (2009). *Evaluating Training Programs*. San Fransisco: Berret-Koehler Publisher Inc. p. 20.
- [7] S. Eko Putro Widoyoko. (2017). *Evaluasi Program Pelatihan*. Yogyakarta: Pustaka Pelajar. Hlm. 173-174.
- [8] S. Eko Putro Widoyoko. (2017). *Evaluasi Program Pelatihan*. Yogyakarta: Pustaka Pelajar. Hlm. 174-175.
- [9] Brikerhoff, RD. Brethower, DM, Hluchyj. T., et.al. (1986). *Program Evaluation a Practitioner's Quide For Trainers and Education*.
Western Michigan: Kluwer-Nijhoff Publishing Lincoln & Arifin Zainal. (2009). *Evaluasi Pembelajaran*. Bandung: Remaja Rosdakarya. Hlm. 182.
- [10] Nuryani Y. Rustaman, et.al. (2003). *Strategi Belajar Mengajar Biologi*. Bandung: UPI, Hlm. 94-95.
- [12] Sugiyono. (2012). *Statistika untuk Penelitian*. Bandung: Alfabeta, Hlm. 61.
- [13] Suharsimi Arikunto. (2006). *Prosedur Penelitian Suatu Pendekatan Praktik*. Jakarta: PT. Asdi Mahasatya, Hlm. 131.
- [14] Zainuddin & Ghodang, H. (2015). *Aplikasi Analisis Jalur: Pendekatan Manajemen Pendidikan*. Bandung: Citapustaka Media, Hlm. 4.
- [15] Meltzer. (2003). *The Relationship Between Mathematics Preparation and Conceptual Learning Gains in Physics: A Possible, Hidden*
- [16] Variable. In *Diagnostic Pretest Scores*. Department of Physics and Astronomy, Iowa State University, Ames, Iowa 50011, *Jurnal Am. J. Physic*. p. 3. Ngalim Purwanto. (2002). *Prinsip-Prinsip dan Teknik Evaluasi Pengajaran*. Bandung: Remaja Rosdakarya, Hlm.102-103.