INTERNATIONAL JOURNA L OF EDUCATION, INFORMATION TECHNOLOGY AND OTHERS (IJEIT)
https://jurnal.unibrah.ac.id/index.php/IJEIT

Vol. 2, No.2, December 2019

IMPLEMENTATION OF KIRKPATRICK MODEL ON TECHNICAL GUIDELINES FOR TEACHER PERFORMANCE ASSESSMENT OF BASIC EDUCATION PERFORMANCE

Purnomo Hadi[1], Basuki Wibawa[2] Moch. Sukardjo[2]

[1]Department of Educational Technology Universitas Negeri Jakarta and Ministry of Education and Culture
[2]Department of Educational Technology Universitas Negeri

Email: purnomo.thesis@gmail.com

Article Info
Article History:
Received: November 22, 2019
Revised: November 30, 2019
Published: December 31, 2019
e-ISSN: 2623-2324
p-ISSN: 2654-2528
DOI: 10.5281/zenodo.3598142

Abstract:

The research method uses an evaluation developed by Kirkpactrik. Data collection techniques using a questionnaire. Data analysis techniques using descriptive analysis. Participant research participants in the Teacher Performance Peenilian Technical Guidance in 2018. Technical Guidance on Teacher Performance Evaluation is carried out with the aim of providing teachers with skills to be able to use the Teacher Performance Assessment instrument to their peers. Evaluations at each important level are carried out to determine the level of effectiveness. Based on the results of the study obtained the following data; (1) the level of reaction of participants stated that the implementation of Bimtek was very good; (2) at the level of learning participants have a very good understanding of the PKG Guidance material. It was concluded that the implementation of the Teacher Performance Evaluation Technical Guidance was carried out in accordance with the implementation guidelines and objectives. The teacher has a very good level of understanding after following the Teacher Performance Assessment Technical Guidance. It is recommended that Technical Guidance on Teacher Performance Evaluation can be carried out in greater numbers so that the ratio between the assessed and the assessor gets smaller.

Keywords: Teacher Performance, Evaluation, Technical Guidance, Technical Guidance Evaluation, Kirkpatrick model

INTRODUCTION

Technical Guidance (Bimtek) Teacher Performance Evaluation (PKG) conducted by the Directorate General of Teachers and Education Personnel has been running for six years. This PKG technical guidance aims to provide teachers with knowledge and skills to become a teacher performance appraiser. The teachers after following the technical guidance are expected to be able to socialize, guide, and assess teacher peers. The teachers who participated in the technical guidance were trained in a 40 hour pattern. So the results of technical guidance can be known after the teacher returns to their respective work units and performs their obligations as an assessor. There is a fairly long period of time from the implementation of technical guidance until the teacher can perform the obligations as an assessor.

Implementation of PKG technical guidance is important for evaluation. Evaluation is carried out to measure the level of success of the planned program. Evaluations can be made while the program is in progress and after the program is finished. Tayibnapis (2013, pp. 36-37) states there are two types of evaluations namely formative evaluation and summative evaluation. Formative evaluations are carried out during the program to provide useful information to students or parents for program improvement. A summative evaluation is carried out at the end of the program to inform students or parents of the final results of the learning program that has been carried out.

Formative and summative evaluation can also be applied to training, technical guidance, workshops, workshops, or the like. Evaluation is carried out to determine the level of success when the program is carried out while the summative evaluation aims to measure the success of the program objectives that have been formulated during planning. Both of these evaluations complement each other.

In the context of educational technology evaluation can be carried out covering four aspects. Prawiladilaga and Siregar (2004, p. 5) state that there are four Educational Technology areas namely; (1) theory and practice; (2) design, development, utilization, management and evaluation; (3) process and source; (4) for learning purposes. The evaluation aspect includes the evaluation. In the context of evaluating PKG technical guidance, it can be done on aspects of the process and management. The process of dealing with learning is done when technical guidance is carried out. Management evaluation is related to governance in implementing technical guidance.

Evaluation principles remain a consideration so that the results obtained are valid and reliable. Tayibnapis (2013, p. 7) states that there are seven principles in evaluating namely; (1) focusing evaluas; (2) design an evaluation; (3) gathering information; (4) analyzing information; (5) reporting evaluation results; (6) managing evaluations; and (7) evaluating evaluations. Teacher assessors based on PKG guidelines can identify the focus of evaluation. At this stage what components will be evaluated against peers. The teacher after identifying is continued by designing the evaluation that will be carried out. The design of this evaluation can be a chart of steps for evaluating peers. The assessor teacher can gather information and analyze the information that has been obtained. The assessor teacher then reports the results of the assessment to those in need such as the principal or supervisor. The teacher can also check and re-evaluate the evaluation that has been done.

Evaluation can also be applied to the PKG Technical Guidance. Evaluations can be carried out starting from the implementation phase of PKG to the implementation phase carried out by the teacher after following technical guidance. An evaluation model that fits the PKG technical guidance characteristics can use Kirkpatrick. Use of this evaluation model has been carried out in training or other technical guidance. Research conducted by Rusman Ruskanda, Zulfiati Syharial, and Atwi Suparman (Online Journal of Incandescent, Aguatus, 2018) with the title "Implementation of the Kirkpatrick Model as a Training Evaluation Model". Researchers used the Kirkpatrick four level method to evaluate training. The research objective is to determine the effectiveness of the training program that has been implemented. The results showed that the training program carried out had effectiveness at every level of reaction, learning, behavior, and the results of the training both for teachers and institutions or *stakeholders*.

An evaluation of the implementation of the PKG Technical Guidance not only has an impact on the organizing agency but is also expected to be able to improve teacher competence to the maximum. The teacher who has followed the Technical Guidance then becomes an assessor teacher who has the task and function to assess the teacher's colleague. The results of the assessment of teacher

assessors can be used as a reference to improve teacher quality. The results of the assessment can be used as a reference for the development and fostering of teacher competencies on an ongoing basis.

Evaluation of the program can be done in whole or only in part of the program. The selection of part or all of the evaluations to be carried out is based on the principle of conducting an evaluation. Practical and efficient principles can be the main consideration for evaluating PKG bimtek program. Arifn (2019, p. 18) states that evaluation must be practical and efficient. Based on these principles the evaluation is carried out at the level of reaction and level of learning. At these two levels, it is carried out simultaneously at the time the technical guidance activities are carried out. Evaluation also has easy and practical properties. Evaluation at the level of reaction and learning carried out when technical guidance is carried out provides convenience and practicality. In other words, both levels are carried out at one time.

The principle of absoluteness and efficiency does not reduce the function of the evaluation, namely to find out the success of the implementation in accordance with the stated objectives. Evaluation at the reaction level and learning level still has significance because success at this stage is an important part of the level of behavior and results. Teachers who rate the level of reaction very well certainly have an impact on good results too. Likewise, evaluations at the level of learning with good results correlate with the level of understanding of technical guidance material as well. So the valuation at the reaction level and learning level still has an important meaning on the success of PKG technical guidance as a whole.

Technical Guidance (Bimtek) Teacher Performance Assessment (PKG) is an activity in which there is learning. The PKG Guidance Program is one of the activities aimed at training teachers at the basic education level in order to be able to assess peer teachers. PKG Bimtek participants came from all over Indonesia. Bimtek PKG has been going on for six years now. Participants after participating in the PKG Bimtek program are expected to be able to socialize with colleagues in their respective regions or regions.

PKG Technical Guidance Program in conducting learning means that evaluation can be done. Tayibnapis (2013, p. 15) as quoting Alkin's statement states that evaluation is a process of convincing decisions, choosing the right information, gathering, and analyzing information so that it can report a summary of the data that is useful for decisions in choosing several alternatives. So the evaluation carried out on the Bimtek program is an effort to collect data scientifically. Based on these data can be used by other parties as a basis for making decisions.

Evaluation on PKG Guidance must be planned in accordance with scientific principles so that the data obtained is valid and reliable. Both of these conditions are important so that decisions taken are appropriate and right on target. This is in line with the statement Purwanto and Suparman (1999, p.9) stating that evaluation is the process of applying scientific procedures to gather valid and reliable information to make decisions about education and training programs. So evaluation is one to gather information in a valid and reliable manner. This data is important because it is the basis for making good and correct decisions. In other words, evaluation cannot be done arbitrarily but must be carried out in accordance with scientific rules.

Hamalik (2007, p.153) states that there are five conditions in evaluating namely; (1) has validity; (2) has reliability; (3) objectivity; (4) efficiency, and (5) usability or practicality. The evaluation that will be carried out on the PKG Bimtek program also has these five elements. Validity and reliability using triangulation. Objectives are carried out in accordance with evaluation rules and do not constitute order evaluations. It is efficient because the evaluation instruments are developed and distributed with low cost and relatively short time, and have the usefulness of providing input to the parties related to the ongoing program of teacher professional development.

Tayibnapis (2013, p. 49) as citing Brian and Davis's statement there are three elements in the evaluation namely; (1) bringing knowledge and hope together; (2) gathering information; and (3) formulating an evaluation plan. Evaluation is carried out to find out between expectations and reality. Evaluation is carried out that planning has been carried out in accordance with expected in accordance with the objectives to be achieved. Evaluation is one way to gather information on the success of the planned program. Evaluations are not necessarily carried out but need to be designed in accordance with the objectives to be achieved. So evaluation requires careful planning so that you can measure what you want to measure. The accuracy of the measurement has an impact on the accuracy in decision making based on the evaluation results obtained.

Decision making in the PKG Bimtek program can be seen from two points of view, namely the assessor teacher and the impact it has. The PKG Guidance Program that gave birth to assessors can be said to be successful if the workplace unit is able to perform its functions and duties as an assessment teacher and is able to socialize PKG guidelines to peer teachers. Hamalik (2007, p. 156) states that measurement is an effort to find out how many things have been owned by students from things that have been taught by teachers. This understanding shows that the measurement is quantitative.

Evaluation can be done by external or internal parties of the institution to be evaluated. Evaluation can also be part of scientific research because in principle the evaluation collects valid and reliable information. Purwanto and Suparman (1999, p. 22) state that the internal evaluator is someone who has knowledge and is involved in the development or implementation of the program, while external evaluators are outsiders who do not join the program. Researchers are outsiders who do not join in the evaluation of program planning.

Guidance and Counseling PKG as well as learning done at school, there are students or in this program called participants. To know the success needs to be evaluated. Popham and Baker (2003, p. 112) state that there are two interrelated things, namely the evaluation of students on their learning achievement and how teachers measure these abilities. Both of these are a unified whole that can not be separated. Evaluation has an important role to measure the achievement of learning objectives. Evaluation tools used by the teacher determine the intended learning objectives. There are two types of tests that can be done by teachers, namely tests and nontest. The test can be in the form of written and oral while non-test can be in the form of protoflio or through worksheets, observations, and the like.

At the Bimtek PKG the evaluation is aimed at participants to measure the level of material knowledge provided. The test given can be a test or non-test. In the study of Dewassa people like Bimtek, evaluation often uses a non-test in the form of a questionnaire or a portfolio of participants. Hamalik (2007, p. 159) states that the evaluation of learning outcomes is the entire measurement (data and information gathering), processing, interpretation, and consideration for making decisions about the level of learning outcomes achieved by students after conducting learning activities in an effort to achieve learning objectives which has been set. In the context of PKG Bimtek evaluation, the measurement is carried out on participants, namely teachers.

Evaluation using both tests and nontests still uses quantitative data. The questionnaire often uses data with a Likert scale. Likewise, the assessment of the portfolio has a standard assessment for each item given a scoring. Woolfolk (2008, p. 411) states that the measurement is quantitative, namely a description of an event or characteristic using numbers. Measurements indicate how much, how often, or how well by giving a rating, ranking, or rating. Likewise, measurements made on the PKG Bimtek program using Likert scale measurements. In this measurement ranges from very good, good, enough, or less. Based on the results of these measurements, a conclusion is drawn.

Purwanto and Suparman (1999, p. 8) citing the statements of Cronbach and Suppes stating that evaluation is a process in which relevant data is collected and transformed into information for decision makers. So the evaluation is a collection of data as a basis for making decisions on programs that have been carried out. The evaluation of the PKG Bimtek program can also be used as a basis for decision making by the authorities. Decision making can be in the form of changes or maintaining training models or the resulting impact.

Davies (in Arikunto, 2010, p. 3) states that evaluation is a simple process to give / assign values to a number of goals, activities, decisions, performance, processes, people, objects, and many others. Evaluation is not only related to learning but all activities can be evaluated. Evaluation also aims to provide or assign values to the object being evaluated. So the evaluation is an activity that must be done so that the object being evaluated can be assigned a value.

Muslich (2007, p.) States that performance appraisal is an appraisal based on observers' observations of activities as they occur, assessments are usually carried out on activities that are appearance. In the context of PKG Bimtek performance evaluation conducted on the performance of teacher evaluators to peers. The results of this performance evaluation can be used as a reference for decision making for follow-up plans. Evaluation is the culmination of a long process of learning that has been done. Evaluation has important meaning in learning. Evaluation has at least two functions, namely to find out weaknesses and strengths and feedback. In the first function is to find out the weaknesses and strengths of the program that is being and is finished on the road. At this stage an evaluation can be carried out throughout the process and after the program ends. In the learning system

often used formative and summative terms. Formative evaluation is carried out in the process while summative evaluation is carried out after the process is complete.

Moore (2005, p. 158) states that evaluation is *a viral part of the instructional process. Evalution must be conducted to determine whether are learning, to gauge the appropriateness of the curriculum for given groups of students, to identify what must be discussed, to ensure proper placement of individual students when a program of instruction, and to make sure that state guidelines for achievement have been met.* The PKG Bimtek evaluation was conducted to find out the success of the planned program based on the objectives. PKG Technical Guidance Evaluation is carried out in two stages, namely the stage at the time of implementation, namely the reaction level and the learning level. The second stage, the evaluation is carried out at the level of behavior and results, after participants return to their respective work units and conduct an assessment with colleagues.

Evaluation models used are tailored to the needs and objectives of the evaluation. There are several evaluation models developed by experts. Evaluation models that are often used include the following.

The Bell System evaluation model, this is used in the largest companies in the United States. This evaluation model is used to measure employee competencies in the company. Wirawan (2011, p. 306) states that the Bell System evaluation model is an evaluation tool used in AT&T and the United States Bell System. This model consists of four levels namely; (1) reaction to output, which is measuring the level of acceptance of the program that has been carried out by the company. At this level a number of components such as content, materials, methods and activities carried out are measured; (2) output capability, which measures what participants expect to know, think, do or produce at the end of the program; (3) application output, which measures what participants know, think, do or produce in the work altar prepared by the program; (4) the benefits of the output, which is to measure how high the organization or institution gets the benefits of the program that has been carried out both in the short and long term.

When examined in more depth the Bell System evaluation model has similarities and similarities with the model developed by Kirkpatrick. The only difference is the use of terminology or diction, but in substance have similarities. The Bell System evaluation model reinforces the use of the Kirkpatrick evaluation model to measure the level of reaction and learning of the technical guidance of Teacher Performance Assessment. This technical guidance program also deals with increasing the capacity and capability of teacher human resources. Teachers with certain levels and classes, namely groups III and IV who can only follow the PKG technical guidance program. Group III is divided into four levels, namely III-A, III-B, III-C, and III-D, as well as group IV divided into groups IV-A, IV-B, IV-C, and IV-D. Only those who have a higher class can only judge the lower class. Group III-B, for example, may only class III-A, and so on. Group III-A may not evaluate group III-B.

The CIRO evaluation model is one of the evaluations that is often used for human resource development. This evaluation aims to measure human resource development programs in an institution or organization can run effectively and efficiently. Wirawan (2011, p. 306) states that the *Context evaluation, Input evaluation, Reaction evaluation* and *Outome evaluation evaluation models,* developed by Wrad, Bird, and Rachman are among the models that are often used to evaluate human resource development programs. In *context evaluation it* measures about three aspects namely the final objective which measures the deficiency of HRD will be eliminated or resolved if there are problems. The intermediate objective is an evaluation that measures changes in employee behavior. Immediate objectives are evaluations that aim to measure employees' knowledge, attitudes and skills. In the *input evaluation* aims to measure the existing sources and determine these sources. Reaction evaluation measures the information obtained from participant reactions to improve the process of developing human resources. In outcome evaluation measures the effect of human resource outcomes.

The CIRO evaluation model if considered is a combination of the CIPP evaluation model and the Kirkpatrick model. Development of an evaluation model that combines the two types of models tailored to the needs. The CIRO model places more emphasis on developing human resources in a sustainable manner.

According to Kirkpatrick and Kirkpatrick (2006, p. 3) several factors that must be considered in the training design are 1) determining training needs, 2) setting objectives, 3) determining training content, 4) selecting training participants, 5) determining the right schedule, 6) choosing the right

facility, 7) choosing the right instructor, 8) choosing and preparing the right audiovisual aids, 9) coordinating the program to interested parties 10) conducting program evaluations.)

Evaluation is an important part in designing a training . Evaluation determines how successful the training is both from *input* processand *output* including how to improve the program based on the training. Kirkpatrick (1998: 16) suggests the effectiveness of a training, namely: 1*) To justify the existence of the training department by showing how it contributes to the organizational's objectives and goals 2) to decide whether to continue or discontinue training programs 3) to gain information on how to improve future training programs.* There are three functions in conducting an evaluation namely first to see whether the program carried out in accordance with the objectives to be achieved. Second, based on the evaluation whether the program needs to be continued or terminated, and thirdly to make improvements both at the time the program is implemented and at the end of the program.

Evaluation models in training programs can use models developed by Alkin. When Alkin was head of the evaluation study center at the University of California, (CSE-UCLA, *University of California Los Angeles Center for Study Evaluation*) he developed a framework that was almost as aspects of the CIPP model. The CSE-UCLA model is an evaluation model that has five dimensions of evaluation (*system assessment, program planning, program implementation, programs improvement, program certification*).

The CSE-UCLA model is an evaluation model that has five

evaluation dimensions, including an *assessment system* that provides information about the state of the system, *programs planning* that helps the selection of certain programs to meet the needs of the program, *program implementation* that prepares information to introduce the program, *program improvement* that provides information about program functions / performance, *program certification* that provides information about the benefits or use of the program.

One of the advantages of the model is thecomponent *program*

implementation that is able to introduce the existence of the program being evaluated. However, this model still has weaknesses that have not been able to show quantitative and sequential calculation results accurately and from the highest to the lowest categories of each evaluation component. Thereforemodel *CSEUCLA* thisneeds to be modified so that it can be used to measure overall training optimization and provide a clear picture to identify problems when training is considered ineffective.

The *Context, Input, Process, Product* Model was developed by Stuefflebeam**.** This model starts with the view that the success of educational programs is influenced by various factors, such as: characteristics of students and the environment, program objectives and equipment used, procedures and mechanisms for implementing the program itself. The evaluation of this model intends to compare the performance of various program dimensions with certain criteria, to finally arrive at a description and *judgment* regarding the strengths and weaknesses of the program evaluated by Stufflebeam (2017) stating that the evaluation objectives are as follows; (1) determination and provision of useful information to assess alternative decisions; (2) helps the *audience* to assess and develop the benefits of educational programs or objects, (3) helps the development of policies and programs.

Based on the analysis of the CIPP model, it can be concluded that the strength of the CIPP model includes; Research using the CIPP model has power. The strengths of this model include; (1) CIPP has a holistic approach to evaluation, aimed at providing a very detailed and broad picture of a project, starting from its context to the time of the implementation process. with the CIPP model, context indicators, inputs, processes and products /*outputs* that can be compared not only to determine whether there are differences in objectives with the actual situation,

but also compared to the specified standards; (2) CIPP has the potential to move in theevaluation areas *formative* and *summative*, so it is equally good in helping to make improvements during the program, as well as providing final information.

Evaluation of the Kirkpatrik model (2016) has a significant contribution to the development of training program evaluation concepts as revealed that: "*There is no doubt that Kirkpatrick's model has made valuable contributions to training evaluation thinking and practice"*. The advantage of the model developed by Kirkpatrik is the simplicity of the process in carrying out the evaluation.

Bates stated (2004: 342) that the kirkatrik model is not a perfect model because it only evaluates the effectiveness of training based on outcomes, some important things that need to be evaluated do not receive attention. Interactions in the training process between factors are very complex and cannot be

generalized just to assess training outcomes. Training evaluation requires an evaluation step which explains the training details so that the evaluation results provide a clear picture.

Regarding the evaluation of the training model, Kirkpatrick (2016: 19) that there are four levels to assess the effectiveness of the program namely level 1 reaction, level 2 learning, level 3 Work behavior and level 4 results are mainly related to organizational or community goals.

Tayibnaspis (2013, pp. 15-16) states that the Brinkerhoff evaluation model consists of three evaluation groups arranged based on the incorporation of the same elements, such as other evaluators, but in their own composition and version. The three groups, namely; (1) *fixed vs emergent evaluation design*; (2) *formative vs summative evaluation*; (3) *experimental and quasi experiment design vs. natural inquiry*.

Design evaluation) *fixed* planned and determined systematically before the evaluation model is presented. This evaluation model is carried out in stages with careful planning starting from planning the evaluation, finding information, in accordance with the objectives to be achieved. Cost fixed design is quite expensive, this is due to factors starting from planning to writing an evaluation report. Model *emergent design evalution.* This evaluation design is made to adapt to the influence and situation that is ongoing and developing. This model data is obtained by conducting hearings, problems that arise, when the program activities take place. The weakness of this model is that it takes a lot of time to gather valid and reliable data. The cost required is also quite large.

model *Formative vs. summative evaluation.* Formative evaluation is used to obtain information that can help improve programs, projects, curriculum, training, workshops. This evaluation is carried out while the program is ongoing. This evaluation serves to correct if there are deficiencies in the implementation of the program. Some experts claim that formative evaluation is the key to the success of programs that have been designed to achieve goals. Summative evaluation is an evaluation conducted at the end of the program. Evaluation aims to determine whether a program can be continued or not based on summative evaluation data.

Anevaluation model *experimentaland quasi experiment design vs. natural inquiry*. Both of these models use the classical methodology. In this model the research subjects were carried out using random sampling or randomly. Both of these evaluation models use hypothesis testing. The purpose of using this evaluation model is to assess the benefits of the new object or strategy being tested. The results of this study can be generalized because the research subjects were conducted using a random system. In this evaluation model known as internal validity and external validity.

The concept of Teacher Performance Assessment / PKG is based on the demands of professionalism. referring to the State Ministerial Regulation PAN and RB Number 16 of 2009 defines the Teacher Performance Assessment is the assessment of each item of the main tasks of the teacher in the context of career development, rank, and position. This assessment is carried out through observation and monitoring. The implementation of the main task of the teacher certainly cannot be separated from the ability of a teacher in the mastery of knowledge, skills, and attitudes in carrying out his duties as a professional educator. Teacher Competency Assessment as an illustration of understanding and awareness of the teaching profession. Teacher competence shows a person's behavior, competence is the capacity to carry out duties as an educator. according to the laws of the Republic of Indonesia No. 14 of 2005 concerning Teachers and Lecturers states that teachers are professional educators with the main task of educating, teaching, guiding, directing, training, assessing and evaluating students in early childhood education through formal education, basic education and secondary education.

The results of implementing Teacher Performance Assessment have two functions, namely; (1) is used as a basis for making plans for the development of Continuous Professional Development (PKB) for teachers as learners; and (2) is used to fulfill teacher credit numbers in promotion and position. Surya (2004: 92) revealed 5 competencies that must be possessed by teachers, including; (1) personal competence, namely the quality of a teacher's personal ability needed to be a good teacher; (2) professional competence, namely the various abilities needed to be able to realize themselves as professional teachers; (3) social competence that is the ability needed by someone to succeed in dealing with others; (4) intellectual competence, namely the mastery of various sciences related to their duties as a teacher; and (5) spiritual competence, namely the quality of faith and devotion as religious people.

Based on the background of the research problem, the following research can be formulated: (1) what is the level of conformity of participants with the guidelines for implementing the Teacher

Performance Assessment Technical Guidance (PKG) at the basic education level ?; (2) what is the level of participant's understanding of the Teacher Performance Assessment (PKG) technical guidance at the level of basic education ?; (3) what is the level of implementation carried out on participants who have followed the Teacher Performance Assessment Technical Guidance (PKG) at the basic education level ?; (4) how to increase the participant's knowledge from the Teacher Performance Assessment (PKG) Technical Guidance training at the level of basic education?

RESEARCH METHODOLOGY

Evaluation research using the Kirkpatrick model has different characteristics. This difference in characteristics causes the research design in the evaluation of the PKG Bimtek program to be seen in the following flowchart.

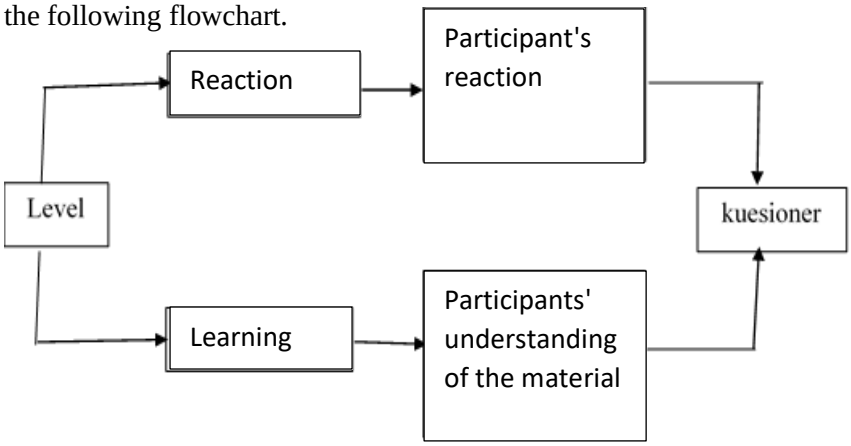

Figure 1 flow of the Bimtek Teacher Performance Assessment

Research data is quantitative data taken using a Likert scale questionnaire. At each level, namely the level of reaction, the level of learning, the level of behavior, and the level of data results obtained using a questionnaire distributed to the sample. The whole analysis process refers to the opinion of Milles and Huberman (1994). Data analysis is known as interactive model analysis. Control and supervision of data needs to be done carefully so that the data obtained is valid and reliable. The steps of data analysis in the interactive model used in this research data analysis are described as follows:

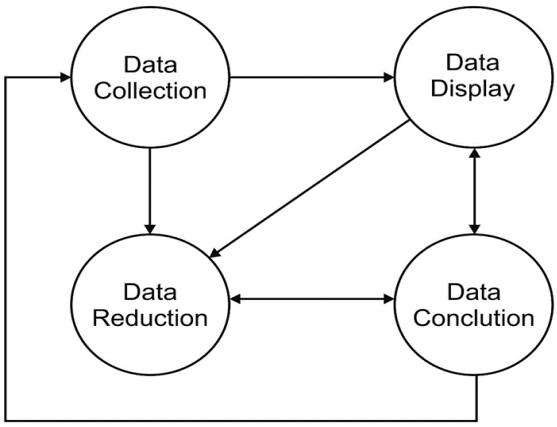

Figure 2. Interactive model of qualitative data analysis (sources: Miles and Huberman, 1992, p. 17.

PKG Technical Guidance is one of the efforts made by the Directorate of Teachers and Education Personnel. Training through Bimtek aims to provide understanding to teachers so that they can become assessors of their peers who have lower class or rank. Evaluation of PKG Bimtek implementation is carried out at the reaction level and learning level. At the reaction level it is divided into three components, namely the evaluation of the implementation of Bimtek, the quality of implementation and instructors, and the assessment and meaningfulness of Bimtek for participants. At the level of learning there are two components that are the focus of evaluation, namely the understanding of technical guidance materials and skills regarding technical guidance materials. Based on the results of the study obtained data as follows.

1. Reaction Level

At the reaction level there are four indicators that are evaluated by participants. In the indicator of the relevance of training there are four sub-indicators. The assessment criteria are score 13-16 = Very Good, 9-12 = Good, 5-8 = Enough, and 1-4 = Poor. Based on the evaluation results at the reaction level with indicators of the relevance of the training obtained research data as follows.

Table 1 Reaction Evaluation Level Reaction indicators indicators of the relevance of training

Category	Respondent	Interval	Percentage
Very Good	21	13 -	70 70%
Good	4	9 -12	13.3%
Fair	3	5 - 8	10.0%
Less	2	1 - 4	6.7%
Total	30		100%

There were 21 participants who said they were very satisfied or 70%, stated that they were Very Good, as many as 4 respondents or 13.3%, participants who stated that they were Good, as many as 3 people or as much as 10%, participants who stated that were Sufficient, and as many as 2 people or as much as 6, 7%. Stating Less Based on these data if a bar chart is made it looks like the following.

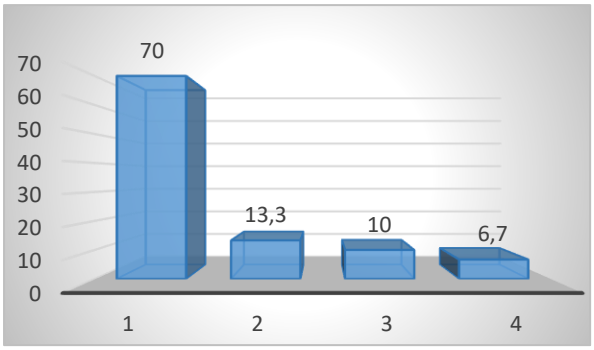

Figure 3. graphs the relevance of the training to the needs of the participants

Evaluations are carried out by participants at the Reaction level on the second indicator namely material and training. The criteria for this indicator are 15 - 20 = Very Good, 10 - 14 = Good, 5-9 = Enough, and 1-4 = Poor. Based on the results of the evaluation conducted by participants can be seen in the following table.

Table 2. Recap Evaluation of Reaction Level Indicator material and training

Category	Respondents	Interval	Percentage
Very Good	20	15 - 20	66.7%
Good	8	10 - 14	26.7%
Fair	2	5 - 9	6.6%
Less	0	1 - 4	0%
Total	30		100%

The mastery of the material by the instructor was considered very good by 20 participants or 66.7%, stated that the material in the Very Good category, as many as 8 participants or 26.7%, stated that the material and training were Good, as many as 2 participants or 6.6% and in Fair category, and in the Poor category, no one judges. Based on the evaluation results on the material and training can be seen in the following diagram.

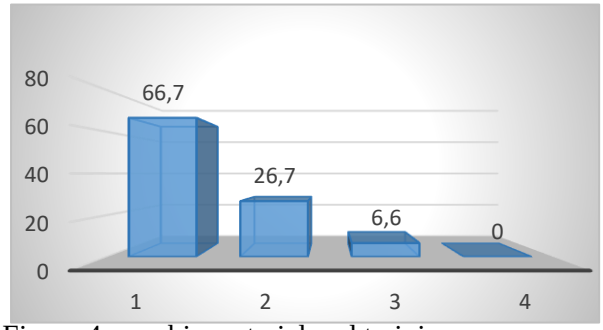

Figure 4. graphic material and training process

Evaluation conducted by participants on the third indicator is the reaction to the instructor or trainer. In this indicator the assessment criteria are set as follows; 19 - 24 = Very Good, 13-18 = Good, 7 - 12 = Enough, and 1- 6 = Poor. Based on the results of the analysis of incoming data obtained data as listed in the following table.

Table 3. Summary Evaluation of Reaction Level Reaction indicators for instructors or trainers

Category	Respondents	Interval	Percentage
Very Good	26	19 - 24	86.7%
Good	4	13 - 18	13.3%
Sufficient	0	7 - 12	0%
Less	0	1-6	0%
Total	30		100%

Based on the results of evaluations conducted by the participants of the instructor on aspects of the ability to manage Bimtek classes obtained data in the very good category provided by 26 participants or as much as 86.7%, in the good category as many as 4 people who gave or as much as 13.3%, and categories enough and less no one gives a rating. Based on these results can be seen in the following graph.

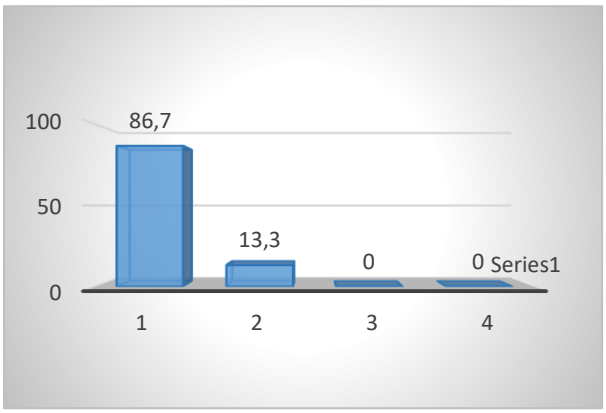

Figure 5. graphs the reaction to the instructor or trainer

Evaluations are conducted by participants at the reaction level with indicators of training facilities and facilities. Evaluations are carried out by participants at the Reaction level on the second indicator namely material and training. The criteria for this indicator are 21 - 28 = Very Good, 14 - 20 = Good, 7 - 13 = Enough, and 1- 6 = Poor. Based on the results of the evaluation conducted by participants can be seen in the following table.

Table 4. Rating of reaction levels of facility and facility indicators

Category	Respondent	Interval	Percentage
Very Good	25	21 - 28	83.3%
Good	3	14 - 20	10%
Sufficient	2	7 - 13	6.7%
Less	0	1-6	0%
Total	30		100%

Based on the results of the evaluation given by participants of the facilities and facilities. As many as 25 participants rated Very Good or 83.3%, assessments in the Good category were 3 participants or 10%, participants who gave Quite ratings were 2 people or 6.7%, and those who gave Less ratings did not exist. Based on the results of these studies can be seen in the following graph.

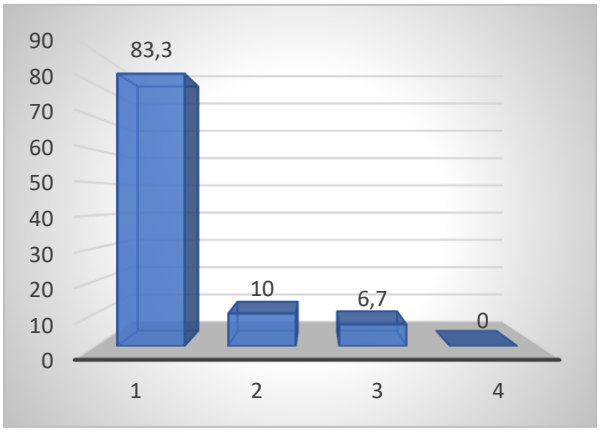

Figure 6. graphic results of PKG participants' achievements

Evaluations at the learning level are focused on increasing participants' understanding of PKG technical guidance material provided by the instructor. Evaluation instruments in the form of tests with an assessment if True is given a value of 1 and if Wrong is given a value of 0. There are 25 items divided into five categories according to the learning outcomes developed by Bloom, namely understanding, application, synthesis, analysis, and evaluation. For knowledge categories united with understanding.

Based on the results of the pretest and posttest conducted during PKG technical guidance the following data was obtained.

Table 4.4 Recap of Learning Level *Pre Test*

Category	Respondents	Interval	Percentages
Very Good	15	19 - 24	50.0%

Good	7	13 - 18	23.3%
Sufficient	6	7 - 12	20%
Less	2	1 - 6	6.7%
Total	30		100%

Based on the data that has been obtained can be explained as follows. There were 15 participants or 50% in the Very Good category, as many as 7 participants or 23.3% in the Good category, as many as 6 participants or 20% in the Fair category, and as many as 2 participants or 6.7% in the Poor category. The results of the pretest can be illustrated through a bar diagram as follows.

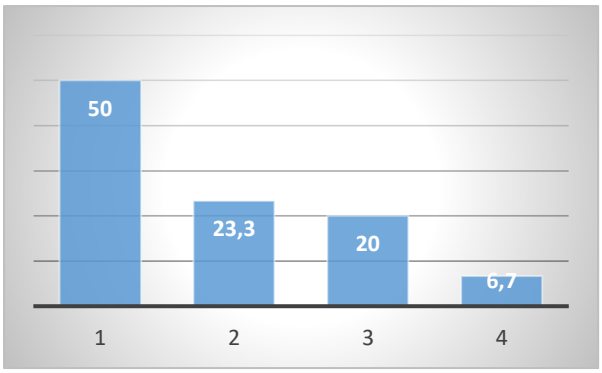

Figure 7. graphic results of PKG participants' achievements

PKG Bimtek participants after doing the pretest then followed the training with a pattern of 40 hours face to face. At the end of the Bimtek post test was carried out using the same test at the time of the achievement. Based on the results obtained by the data can be seen in the following table.

Table 4.4 Recap Level ofLearning *Post Test*

Category	Respondents	Interval	Percentage
Very Good	25	19 - 25	83.3%
Good	3	13 - 18	10%
Sufficient	2	7 - 12	6.7%
Less	0	1-6	0%
Total	30		100%

25 or 83.3% participants were stated in the Very Good category, as many as 3 participants or 10% were stated in the Good category, as many as 2 participants or 6.7% were stated in the Fair category, and in the Less category it was stated as 0%. The results of this post test can be illustrated through the following bar diagram.

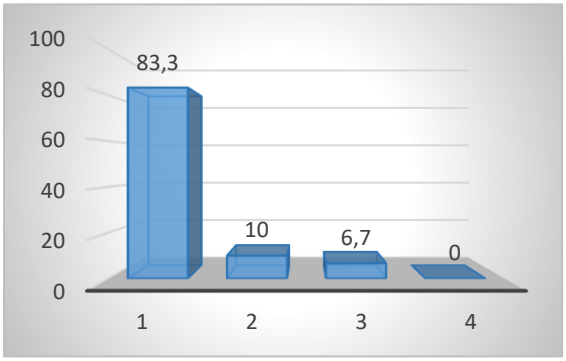

Figure 8. graphs of the results of PKG participants' posttests

Teacher Performance Appraisals conducted periodically for teachers at every level of education are one of the efforts to improve teacher competence in an equitable manner. The Implementation of PKG Technical Guidance is the mandate of the Law, namely Law Number 14 of 2005 About Teachers and Lecturers that the teacher is a professional educator with the main task of educating, teaching, guiding, directing, training, assessing, and evaluating students in the education of children. early formal education pathways, basic education and secondary education. The quality of teaching is one form of work ability that is supported by competence and work motivation as educators and teachers. The teacher has a function as a manager who guarantees the implementation of a quality education process and plays a strategic role in realizing the quality of the school. A quality learning process requires the presence of a professional teacher.

Based on Law No. 14 of 2005 concerning teachers and lecturers, each teacher is required to have a minimum academic qualification of bachelor (S-1) or diploma four (D-4) in accordance with Law No. 14/2005, article 9. Professional teachers must possess minimum competency (pedagogic, professional, social and personal), have an educator certificate, be physically and mentally healthy, and have the ability to realize national education goals. The need for the availability of teachers who have good competence for classroom action research, teaching performance is quite high. The demand for professionalism is not in line with expectations.

To ensure the implementation and delivery of quality education, competency testing of teachers is carried out in practice and in research. Competency test is part of the teacher performance appraisal system to ensure the implementation of a quality education system. In accordance with the policy to realize an adequate teacher performance evaluation is Permenpan no. 16 of 2009 and Ministry of Education Regulation no. 35 of 2010. Conceptual assessment provides the basis for improvements in teacher performance both for teaching and research. The results of the competency test indicate the value achieved is not in accordance with the target or adequacy standards for the delivery of adequate education.

The practice of organizing teacher competency assessments as an instrument for improving teacher performance is still functioning administratively. The practical function of providing education as a form of communication between teachers and the world of education is not yet optimal. The benefits and objectives of PKG to provide feedback based on observations on teacher performance have not had much impact on changes in teacher competence, including in West Java. This can be seen from the follow-up results of the teacher's performance evaluation with the reality in the field. There is a gap between the results of teacher professionalism assessment and the reality on the ground. Many teachers do not understand the substance of performance appraisal.

At the PAUD level, teachers who do not have Diploma IV or undergraduate education are ranked first at 32%, while the smallest is at the SMA level, 2.8% of teachers who have less than Diploma IV or undergraduate level. At the junior high school level, 7.1% did not have a Diploma IV or bachelor level of education, while 92.9% of teachers had a Diploma IV or bachelor level or equivalent. At the SMK level as many as 6% do not yet have a Diploma IV or undergraduate level of education, while 94% have a Diploma IV or undergraduate level and equivalent. In SLB, as many as 11% do not have an equivalent level of Diploma IV or undergraduate, while as many as 89% have an equivalent level of Diploma IV or undergraduate and its equivalent.

Based on the data it is known that the number of teachers who do not yet have certification is very large. 56.74% of elementary school teachers are not yet certified. 52.5% at the junior high level, 52.6% at the level. Vocational Teachers is 66.1%. the highest is PAUD teachers by 74.1% of teachers not yet having certification. This shows the guarantee of teacher performance in accordance with the demands of quality is less guaranteed. At the elementary level of education both in elementary school (SD) teachers and junior high schools (SMP) in quantitative terms only 50% are certified and declared as professional teachers. The professional category is obtained through the Teacher Performance Assessment stage. Ideally, every level of education for both teachers in public and private schools is 100% professional and regular Teacher Performance Assessments have been conducted.

The benefits and objectives of PKG to provide feedback based on observations on teacher performance have not had much impact on changes in teacher competence. This can be seen from the follow-up based on the results of the teacher's performance evaluation with the reality in the field. There is a gap between the results of teacher performance assessments and the reality on the ground. Many teachers do not understand the substance of the use of teacher performance instruments, including the development of practices based on relevant theories.

An important aspect in the process of evaluating 14 teacher performance competencies is the appraiser's understanding of assessment instruments. Appraiser's knowledge of 14 competencies is not only limited to memorizing the 14 competencies, but to the practical and conceptual evaluation stage. The assessor has the best theoretical knowledge and is able to explain how the practice of the 14 components.

In addition to knowledge of 14 competencies, procedural and factual knowledge about teacher performance should be mastered. The teacher's ability to assess teacher competency basically determines how successful the PKG process flow mechanism is. Appraiser's understanding of the guidelines, competency assessment instruments and competency indicators as well as how the three guidelines are implemented can produce knowledge in accordance with the facts of the assessment.

The successful implementation of teacher performance assessments has been carried out since 2013 (6 years). Implementation of PKG has not been carried out properly. Some of the problems start from the most common things: socialization of PKG assessment in the field, a policy system to guarantee and support PKG especially in areas that have low Human Resources (HR) support to conduct PKG, supporting sarpras, regional policies.

The results of observations on the performance of teacher behaviors who have followed PKG obtained some description that there are problems in PKG assessment in West Java, especially in Depok. Observations by assessors are not comprehensive. Assessment activities are limited in the classroom. The fact of knowledge that is the basis for conducting an assessment is not yet fully supported by document studies, discussions, learning / coaching processes, colleague interviews, interviews with students, parents. The source of knowledge on which the assessment is based is limited to individual observation (individual epistemology).

The ratio of teachers who become assessors and those assessed is considered as one of the factors in assessing teacher performance. The results of teacher performance appraisal can be used as a basis for consideration of decision-making related to promotion, promotion, the basis for giving a teacher a visit, as well as matters relating to teacher competence and service fees received. The Ministry of Education and Culture annually tries to reduce the ratio between appraisers and those assessed. The Technical Guidance Program (Bimtek) Teacher Performance Evaluation (PKG) is an effort so that teachers have the ability to assess themselves and their peers. The results of this program are expected by the teacher after attending the Technical Guidance to be able to conduct performance assessments for themselves and also colleagues. The effectiveness of the PKG Bimtek implementation that has been done many times has never been evaluated. Based on the results of the evaluation carried out improvements for those who are still less than optimal and improved and maintained for the good. The results of the evaluation in addition to obtaining data on the level of success with the objectives to be achieved, but also the overall impact of training.

Evaluation using steps developed by Kirkpactrick Evaluation was carried out covering four stages. Each stage has a correlation with the next stage. Research on the importance of evaluating the program that has been carried out was raised by Rusman Ruskanda, Zulfiati Syharial, and Atwi Suparman (Online Journal of Incandescent, Aguatus, 2018) with the title "Implementation of the Kirkpatrick Model as a Training Evaluation Model". Researchers used the Kirkpatrick four level

method to evaluate training. The research objective is to determine the effectiveness of the training program that has been implemented. The results showed that the training program carried out had effectiveness at every level of reaction, learning, behavior, and the impact of the training both for teachers and institutions or *stakeholders*.

The results of the study, especially at the level of learning, if noted there is a very significant spike in the understanding of the material by the participants. This condition can be explained as follows.

First, PKG technical guidance aims to enable participants to use assessment instruments properly and correctly. The optimal and maximum mastery of technical guidance material is one of the objectives of this technical guidance. Thus, the material is designed to be understood by participants during the bimtek which is only 40 hours face to face. Participants must be able to use instruments, assess peer teachers, make reports, and verify assessed teachers. This competence is reflected in the pre-test and post-test questions.

Second, technical guidance focuses on the competencies of participants so that the material is designed using case studies that might be found by the participants as assessors. This material is combined with a simulation method so that participants are given a case study to be solved in accordance with the case they face. These materials and methods are considered effective so that participants can understand the use of assessment instruments more comprehensively. This is one way to increase the ability of participants more effectively and efficiently.

Third, the instructor has experience in using adequate learning methods and strategies so that they are easily understood by the participants. Group collaboration, simulations, using worksheets as if assessing peers, is an appropriate strategy implemented with a limited number of face-to-face hours.

CONCLUSIONS AND RECOMMENDATIONS

Based on the results of the study can be concluded as follows; (1) at the reaction level the overall data obtained by the participants stated that the organizer was very good at carrying out the PKG technical guidance. Thus it can be concluded that the organizers have successfully carried out PKG technical guidance activities in accordance with the expected goals .; (2) at the level of learning there is a significant increase in understanding of the material with differences in the results of the pre test and post test. Based on these results it can be concluded that there is a significant increase in the understanding of technical guidance participants.

Based on the results of research on the evaluation of PKG technical guidance it can be recommended to several parties, both institutions and individuals, viz. (1) PKG technical guidance participant teachers should conduct socialization to peer groups so that they can know and understand the assessment grid early on; (2) teachers participating in PKG technical guidance should be able to provide guidance to teachers who are assessed so that they can meet the requirements needed at the time of the assessment; (3) The Directorate General of Teachers and Education Personnel is bound to map out the distribution of teachers who have received technical guidance so that it is not concentrated in one city. Teachers from remote and outermost areas can be included so that they can become assessors for their peers; (4) The organizing team should be able to improve services to participants so that it makes it easier for participants to be able to understand technical guidance material more effectively and efficiently.

REFERENCES

Arifin, Zaenal (2019) *Evaluasi Program: Teori dan Praktik dalam Konteks Pendidikan dan Nonpendidikan*, Bandung, Remaja Rosdakarya.

Basuki, H. (2006). *Penelitian Kualitatif Untuk Ilmu-Ilmu Kemanusiaan Dan Budaya*. Jakarta.

Buckley,R. , dan Caple,J. (2007). The theory and practice of training.London Kogan Page

Danim Sudarwan. 2002. *Pengembangan Profesi Guru*. Jakarta: kencana Pernada Media Grou

Gilley.,J dan Eggland,S (1989) *Principles of human resource development*. A Maycunich: Basic Books.

Hamalik, Oemar (2007) Kurikulum dan Pembelajaran, Jakarta, Bumi Aksara.

Kirkpatrick, D. L. (1998). *Evaluating Training Programs, Second Edition, The Four Levels, San Fancisco*: Berrett-Koehler Publishers, Inc.

Kirkpatrik, D. L. & Kirkpatrick, J. D. (2006). *Implementing the four levels. Practical guide for effective evaluation of training programs*. California: Berrett-Koehler Publishers

Muijs, Danile, David Reynolds (2008) *Effective Teaching: Teori dan Aplikasi*, penterjemah Helly Prajitno Soetjipto dan Sri Mulyantini Soetjipto, Yogyakarta, Pustaka pelajar

Muslich, Masnur (2007) *Pembelajaran Berbasis Kompetensi dan Kontekstual*, Jakarta, Bumi Aksara.

Moore, Kennethd D. (2005) *Effective Instructional Strategies*, California, Sage Publication Inc.

Prawiladilaga, Dewi Salma & Eveline Siregar, ed. (2004) Mozaik Teknologi Pendidikan, Jakarta, Kencana dan Universitas Negeri Jakarta.

Puwanto, Atwi Suparman (1999) *Evaluasi Program Diklat*, Jakarta, STIA LAN Press.

Popham, W James, dan Eva L. Baker (2003) *Teknik Mengajar Secara Sistematis*, penerjemah Amirul Hadi, Jakarta, Rineka Cipta.

Sardiman (2008) *Interaksi dan Motivasi Belajar Mengajar*, Jakarta, RajaGrafindo Persada.

Stufflebeam, D. L., & Zhang, G. (2017). *The CIPP evaluation model: How to evaluate for improvement and accountability.* Boston, MA: Klewer Nijhoff.

Sudjana D. (2007). *Manajemen Program Pendidikan* : *untuk Pendidikan Nonformal dan Pengembangan Sumber Daya Manusia*, Bandung : Falah.

Suparlan, (2006) *Menjadi Guru Efektif.* Yogyakarta.Hikayat

Surya, Muhammad. (2004). *Psikologi Pembelajaran dan Pengajaran.* Bandung: Yayasan Bhakti Winaya.

Syah, Muhibbin. (2010). *Psikologi Pendidikan dengan Pendekatan Baru*. Bandung: PT Remaja Rosdakarya.

Tayibnapis, Farida Yusuf (2013) *Evaluasi Program dan Instrumen Evaluasi untuk Program Pendidikan dan Penelitian*, Jakarta, Rineka Cipta.

Taylor-powell, E., Steele.S., Douglas.,M (1996). *Planning a Program Evaluation* .Wisconsin. University of Wiconsin-extension

Vleuten,C.P.M V.d & Schuwirth,L.W.T (2005) *Assessing professional Competence: from Methods to Programmes*. University of Maastricht, Department of Educational Development and Research.

Yusuf, Muri A. (2015) *Asesmen dan Evaluasi Pendidikan: Pilar Penyedia Informasi dan Kegiatan Pengendalian Mutu Pendidikan*, Jakarta, Kencana.

Wirawan (2011) *Evaluasi : Teori, Model, Metodologi, Standar, Aplikasi, dan Profesi*, Jakarta, RajaGrafindo Persada.

Woolfolk, Anita (2008) *Education Psychology Active Kearning*, penerjemah Helly Prajitno Soetjipto dan Sri Mulyanti Soetjipto, Yogyakarta, Pustaka Pelajar.

Resources from journal

Abdulghani, H., Khamis, N., Majeed, A., & Drees, A. Al. (2014). *Research methodology workshops evaluation using the Kirkpatrick ' s model : Translating theory into practice*, (April).

Aryadoust, V. (2016). Adapting levels 1 and 2 of Kirkpatrick's model of training evaluation to examine the effectiveness of a tertiary-level writing course. *Pedagogies: An International Journal, 12*(2), 151-179

Bates, R., & Bates, R. (2004). A Critical Analysis of Evaluation Practice : the Kirkpatrick Model and the Principle of Beneficence A critical analysis of evaluation practice : the Kirkpatrick model and the principle of beneficence, *Evaluation and Program Planning* 27 (3), 341-347

Bates, R. (2016). A Critical Analysis of Evaluation Practice : the Kirkpatrick Model and the Principle of Beneficence A critical analysis of evaluation practice : the Kirkpatrick model and the principle of beneficence, (August 2004). https://doi.org/10.1016/j.evalprogplan.2004.04.01

Evans, L. (2008) *Professionalism, professionality and the development of education professionals.* British Journal of Educational Studies, 56 (1). pp. 20-38

Harden, A., Peersman, G., Oliver, S., Mauthner, M., & Oakley, A. (1999). *A systematic review of the effectiveness of health promotion interventions in the workplace, 49(8), 540–548.*

Holloway, J., & Holloway, J. (2018). Teacher evaluation as an onto-epistemic framework Teacher evaluation as an onto-epistemic framework. *British Journal of Sociology of Education*, *0*(0), 1–17.

Horn, K.-P. (2016). *Profession, professionalisation, professionality, professionalism — historical and systematic remarks using the example of German teacher education. British Journal of Religious Education, 38(2), 130–140.*Malki et al. 2003)

Matthew A. Kraft1 and Allison F. Gilmour (2016) Can *Principals Promot Teacher Development as Evaluators? A Case Study of Principals' Views and Experiences*

Musal B, Taskiran C, Gursel Y, Ozan S, Timbil S, Velipasaoglu S. 2008. *An example of program evaluation project in undergraduate medical education.* Edu Health 21(1):1–7.

Nazir, M. (2003). *Metode Penelitian*. Jakarta: Ghalia Indonesia Noe, et.al (2011: 106

Noe, R.A., Hollenbeck, J.R., Gerhart, B.,Wright., P.M (2010).*Manajemen Sumber Daya Manusia*. Jakarta: Salemba empat

Olondo Wilfrida Arnodah, (2013),"Teacher improvement through peer teacher evaluation in Kenyan schools", *European Journal of Training and Development, Vol. 37 Iss 7 pp. 635 – 645*

Praslova, L. (2010). *Adaptation of Kirkpatrick's four level model of training criteria to assessment of learning outcomes and program evaluation in Higher Education*. Educ Asse Eval Acc 22:215–225.

Reio, T. G., & Smith, D. H. (2017). A Critique of Kirkpatrick ' s Evaluation Model. New Horizons in Adult Education & Human Resource Development *29*(2), 35-53

Saiders,R (2012) *Assessment of professional development for teachers in the* vocational education and training sector: An examination of the Concerns Based Adoption Model.A *ustralian Journal of Education,Vol. 56, No. 2, 2012, 182–204*

Setaro, J. (2001). *Many happy returns: Calculating e-learning ROI.* Retrieved December 6, 2004, from (http://www.learningcircuits.org/2001/jun2001/Elearn.ht)

Steinert, Y. Mann.K, Centeno.A,, Dolmans,D., Spencer,j.,, Gelula,D., & Prideaux.,D. (2006) *A Systematic Review of Faculty Development Initiatives Designed to Improve Teaching Effectiveness in Medical Education*: BEME Guide no. 8. Medical Teacher, vol. 28, no. 6, pp. 497–526

Vleuten,C.P.M V.d & Schuwirth,L.W.T (2005) *Assessing professional competence: from methods to programmes. University of Maastricht, Department of Educational Development and Research*.