# AN EVALUATION MODEL OF
# ISLAMIC LEARNING EDUCATION PROGRAM IN MADRASAH ALIYAH

*Anidi*
STKIP Pelita Nusantara Buton
Jl. Pahlawan Km. 4 Baubau. Kota. Kota Baubau - Prop. Sulawesi Tenggara, 93716 Indonesia
Email: said_anidi@yahoo.com

**Abstract**

The purpose of this study is: (1) to create the products as the evaluation model learning programs of Islamic education in madrasah aliyah, suitable, precise, and accurate, (2) to know how to test the suitability of the model to obtain the evaluation model of Islamic Education learning programs in madrasah aliyah which is valid, and reliable, and (3) determine the effectiveness of the application of evaluation model of Islamic learning programs that was developed. This was a research and development method study (R & D), referring to the model of Borg & Gall, the development process was simplified into three steps, namely: (1) the stage predevelopment of models, (2) the stage of development of the model, and (3) the operational phase of the field (test model). The data were collected using questionnaires, interview, observation, documentation, and test. Data were analyzed using quantitative and qualitative. The conclusion of this research and development is: (1) evaluation model of Islamic Religious Education Learning program developed comprising: (a) the evaluation procedures, (b) the evaluation instrument, and (c) the evaluation guide. (2) according to the experts, PAI teachers, and principals, procedures, instruments, and evaluation guidelines developed are already good, and can be used, d) the instrument developed is entirely valid (load factor > 0.3), reliable (CR > 0.7), and qualified as a fit model (RMSEA ≤ 0.08 and NFI, CFI and GFI > .90). second, the expert's opinion, PAI teachers and principals, the developed evaluation model (EPH model) is declared effective, practical, and easy to use, and supported by a valid and reliable instrument.

**Keywords:** *model of evaluation, evaluation process, evaluation outcome, Islamic learning education*

## Introduction

People's lives are faced with the issue of decline in manners and ethics in social life, both in the family environment, school (madrasah), as well as in the surrounding environment. The phenomenon of public life today show the stronger need to optimize the implementation of Islamic Education in Madrasah, so able to become the basis of thinking and behave in life.

Religious Education today have many kinds of sharp criticism, because inability to cope the important issues in the life of societies (Sutrisno, 2006, p.5). The facts showed that the influence of non-educative foreign cultures global as well, such as the culture of materialism, consumerism and hedonism which is getting stronger (Muhaimin, 2009, p.51). Religion Education should get a common concern in building the character and moral of the nation.

Ideally, Islamic Religious Education (PAI) able to be the basis of other education, as well as being one of the measures in building national character (Nation character building) and the personality of the learner is balanced both in intelligence quotient (IQ), emotional quotient (EQ) and spiritual quotient (SQ) (Agustian, 2003, p. 175). More pronounced in Law Number 20 Year 2003 on National Education System, that the implementation of PAI in madrasas includes planning, implementation and assessment systems should be well planned, integrated, and sustainable with regard to all aspects of the student whether cognitive, affective and psychomotor.

Various efforts have been made by the government to improve the quality of National Education, namely through the development and improvement of curriculum, improvement of evaluation systems, development of learning materials, procurement of books and tools of learning, improvement of education infrastructures, increased the teacher kompentence, as well as improving the quality of school leadership. As a formal educational institutions, implementation of education in Madrasah should refer to the National Education Standards.

Islamic Religious Education Learning at the school is dependen on factors educators or teachers. Teachee is one component that has an important role Because in addition to as exemplary figure and teacher also as a facilitator, administrator, motivator, counselor, organizer, and Evaluators. The role of the teacher as a manager of learning (learning manager) requires teachers to manage and create a climate conducive learning and fun in order to form the religious competence of students as a whole. Additionally, the learder and Teachers PAI demanded to be resilient and creative in order to create a conducive learning environment Islamic Education.

The creation of a religious atmosphere in madrasah namely by utilizing a mosque or prayer room to pray and practice activities of worship, selebration Islamic great days, the religious broadcasting, convey the values of religious and moral messages to all subjects, utilizing the religion figures in the communities as a learning resource, and the availability of religious laboratory. All of them would help to achieve the learning objectives of the maxsimal Islamic education.

Improving the quality of education be conducted continuously by improving the quality of learning in all fields of study, so that the education goals can be achieved effectively and more efficient. Focus or direction in improving the quality of education is achieved the educational goals as the ability of an intact self-learners, which include the ability of academic or intellectual capital, and moral capacity or moral capital (Zamroni 2005, p. 1). Both the authorized capital is the capital required to improve the quality of education.

Duties and responsibilities of Islamic Education in Madrasah Aliyah not only on religious teacher, but also the responsibility of the madrassa as a whole. Madrasah environment should support and be a laboratory for the study of Islamic education. Improving the quality of learning requires efforts enhancement the whole evaluation learning program because the nature of the learning quality is the quality of the learning

implementation program that has been designed.

Improving the quality ofthe program implementation requires enhancement in aspects of other programs, such as program designing and program components. Improving the quality of the learning program components includes: learning objectives, teachers, students, materials, sources of learning materials, as well as the learning facilities.

Based on the results of preliminary research conducted at several research sites which concluded that: (1) input as input in learning, namely learning completeness inadequate facilities, especially the subjects of jurisprudence requiring laboratory practice; (2) the learning process associated with the performance of teachers in the classroom, the teacher is still dominated learning; and (3) assessment of learning outcomes they prefer the cognitive, while affective and psychomotor were the less noticed.

Research conducted Schneider (Morrison, Mokasi, & Cotter, 2006, p. 5) concluded that the physical environment of the classroom or learning facilities had a significant impact on student learning and teacher performance. Classrooms were uncomfortable, hot, cold and many people were passing so as an obstacle to achieve the better learning. the next Results of research conducted by Sudjana (2002, p. 42) showed that 76.6% of student learning outcomes were influenced by the performance of teachers, it can be seen the ability of teachers to teach contributed 32.43%, mastery of the subject matter contributed 32, 38 % and the attitude of teachers to the subjects contributed 8.60%. While research conducted Wahyudi (2003, p. 1) proveded that there was strong correlation between student achievement in a class with moods or social environment created in the class.

Based on preliminary studies and empirical data mentioned above, to more optimize the results of the evaluation of the Islamic learning Education in Madrasah Aliyah, an evaluation of the learning program needs to be made more comprehen-sive, with scope of program not only the aspect of learning outcomes, but also included the aspect of the learning process.

Ideally Madrasah Aliyah according to Saleh, (2004, p.47) produced a profile that describe graduations as follows: (1) had the faith and piety in accordance with the teachings of his religion, (2) had a base value humanioran to implement a unity in life, (3 ) master academic ability and skills as well as the ethos of learning to continue education, (4) Converting academic ability and life skills in local and global communities, (5) the ability of expression, appreciated art and beauty, (6) has the ability to exercise, maintaining healthy, build endurance and physical fitness, (7) participated and nationality insight into the life of society, nation and state as a democracy.

Islamic Education in Madrasah Aliyah needs colaborate between knowing (knowing), doing (practice), and being (live), so that the necessary adaptation learning program evaluation that is implemented. Mardapi (2000, p. 2) explained that the evaluation in the field of education in review of the target can be divided in two parts, namely the evaluation of macro and micro evaluation. Evaluation of macro targets are education programs in general, the program planned to improve education. Evaluation of micro frequently used at the classroom level. So the micro evaluation targets are learning in the classroom program.

Evaluation of the learning program is more important is how the learning of Islamic Education could be evaluated in a professional manner, so provided the accurate and comprehensive information. It required a program evaluation model specifically developed to evaluate the components of the learning program that is implemented.

In order to achieve successful learning programs PAI in madrasa, requred the suitable evaluation model so could provide the accurate information to stakeholders and more important to the parties concerned, especially the leadership of madrasas, both in terms of content, scope and format of evaluation to improve the learning program.

This study developed an evaluation model of Islamic Religious Education Learning program (PAI) at Madrasah Aliyah. Evaluation model developed was limited to components: (1) the evaluation procedure (2) the evaluation instrument, (3) the evaluation guide. Therefore, the problem would be examined in this study, as follows: (1) how the evaluation model of learning programs PAI to provide the comprehensive information, axactly and accurate, and could contribute the benefit to the leadership and the teacher of Madrasah?, (2) how the suitability evaluation model of learning program PAI at Madrasah Aliyah valid and reliable? and (3) how the effectiveness of the learning program evaluation model developed?

Based on the formulation of the problem mentioned above, the purpose of research and development are as follows: (1) develop a model evaluation of learning programs that provide comprehensive information, exactly and accurate, (2) find out the suitability evaluation model of learning programs are valid and reliable, and (3 ) find out  the effectiveness evaluation model of learning programs developed.

**Research Methods**

This research is development research (research and development), which aims to produce a product in the form of a model of evaluation, namely the learning evaluation program of Islamic education in Madrasah Aliyah. The research model used based the models of the research and development by Borg & Gall (1993, pp. 275-276), which consists of ten steps: (1) research and information colecting, (2) planning, (3) the first product development , (4) the primary trials, (5) the revision of major products, (6) the trial primary, (7) the revision of operational product, (8) trial field operations, (9) the revision of the product, and (10) the dissemination and implementation.

Ten steps which refer to the development of research models Borg & Gall to adjust into three steps, namely: (1) pre-development stage models, (2) the stage of the model development, and (3) the stage field operational (test model). Pre-development stage models conducted with observations, interviews, and documentation by the teachers and the headmaster as well as reviewing relevant literature and research.

The development phase was done by determining the model of program evaluation with validation instrument colleagues (first draft), validation expert (expert judgment), and the drafting II. The field operational phase (test model) was done with Phase I trial (legibility), Phase II trial (feasibility), and Phase III trial (field operations).

The Subjects in this study consists of the students, teachers and headmaster. the sampling technique was done by purposive random sampling, namely the classified of the good quality, medium, and low. Madrasah was used as test subjects were 15 madrassas, which was divided into five good quality madrasas, 5 medium quality madrasa, and 5 Madrasa are low quality. Madrasah is spread several Ministry of Religious Regency/City region of Southeast Sulawesi province.

Data in this research are quantitative and qualitative data. Data collection techniques by used interviews, observations, questionnaires, and documentation. The research instruments was the assessment sheet, interview guidelines, observation guidelines, questionnaires, documentation guidelines, and sheets of observations, and tests.

Validity of the instrument in this study included content validity (content validity), and construct validity. The content validity obtained from the judging of the Experts and teacher of PAI towards the research instruments that had been prepared by using the formula Aiken's V. Testing of validity and reliability construct consist of the feasibility of the instrument by using using the correlation product moment and Alpha Cronbach, whereas the suitability model of the instrument used confirmatory factor analysis (CFA) by using  lisrel program 8.70.

The parameters was used for testing the validity namely used the product mo-

ment correlation whereas reliability test used Cronbach Alpha. Validity viwed from coefficient r> 0.3 (Norušis, 1986, p. 12). To establish the reliability of the constructs, used Cronbach's Alpha formula is at least 0.7. While testing the suitability of theoretical models with empirical data evaluation model in this study refers to the criteria of Goodness of Fit (GOF) is the Root Mean Square Error of Approximation (RMSEA) <0:08; Chi-Square were obtained from the test has a probability greater than 0.05 (p> 0.05), and Normet Fit Index (NFI), Comparative Fit Index (CFI), and Googness of Fit Index (GFI)> 0.90.

**Finding and Discussion**

Pre-development Stage

The results of observation, interviews, and data documentation on pre-development stage became the base as need assessment for developing model evaluation program learning of Islamic education in Madrasah Aliyah. Results of studies development modal evaluation lerning program then combined with the results of the literatures related to the program evaluation model, then translated in initial model. This initial product was developed of the model CIPP that has two stages of evaluation namely a process, and learning outcomes, hereinafter

the evaluation model program named EPH models, as presented in Figure 1.

Product Development Phase

Based on the results of the pre-development studies then performed subsequent drafting of the prototype model evaluation development of Islamic Religious Education Learning program at Madrasah Aliyah. The prototype consists of: (1) components and procedures learning program evaluation model of Islamic education in Madrasah Aliyah, (2) the istruments of learning process, and learning outcomes, and (3) the using of manual evaluation model learning programs. The products had been designed as a first prototype that was developed through self evaluation validated by experts of Islamic education, and expert PEP, (expert review) based on the contents, and language. The validation method used the peers instrument validation and the Delphi technique.

Experts and Teachers PAI provided feedback and agreed with the evaluation procedures components, instruments, and evaluation guidelines. The results of validation model evaluation program developed, the item's score used a calculation coefficient item Aiken's V, as presented in Table 1.

Table 1. Summary Of The Result Validation model Evaluation (Model EPH)

| Evaluation Procedure | | Learning Process | | Learning Outcomes | | Evaluation Guidlines | |
|---|---|---|---|---|---|---|---|
| Teacher | Expert | Teacher | Expert | Teacher | Expert | Teacher | Expert |
| 0.88 | 0.79 | 0.79 | 0.88 | 0.75 | 0.88 | 0.88 | 0.88 |
| 0.83 | 0.75 | 0.75 | 0.79 | 0.75 | 0.79 | 0.71 | 0.79 |
| 0.75 | 0.75 | 0.75 | 0.83 | 0.75 | 0.79 | 0.75 | 0.83 |
| 0.83 | 0.83 | 0.75 | 0.75 | 0.75 | 0.75 | 0.79 | 0.79 |
| 0.75 | 0.75 | 0.79 | 0.83 | 0.75 | 0.83 | 0.79 | 0.83 |
| 0.83 | 0.83 | 0.88 | 0.88 | 0.79 | 0.79 | 0.79 | 0.83 |
| 0.75 | 0.75 | 0.88 | 0.88 | 0.71 | 0.75 | 0.71 | 0.75 |
| 0.75 | 0.75 | 0.88 | 0.88 | 0.75 | 0.75 | 0.71 | 0.75 |
| 0.75 | 0.71 | 0.71 | 0.71 | 0.75 | 0.75 | 0.71 | 0.79 |
| | | 0.75 | 0.75 | 0.83 | 0.83 | | 0.75 |
| | | 0.75 | 0.75 | | | | |
| | | 0.79 | 0.83 | | | | |
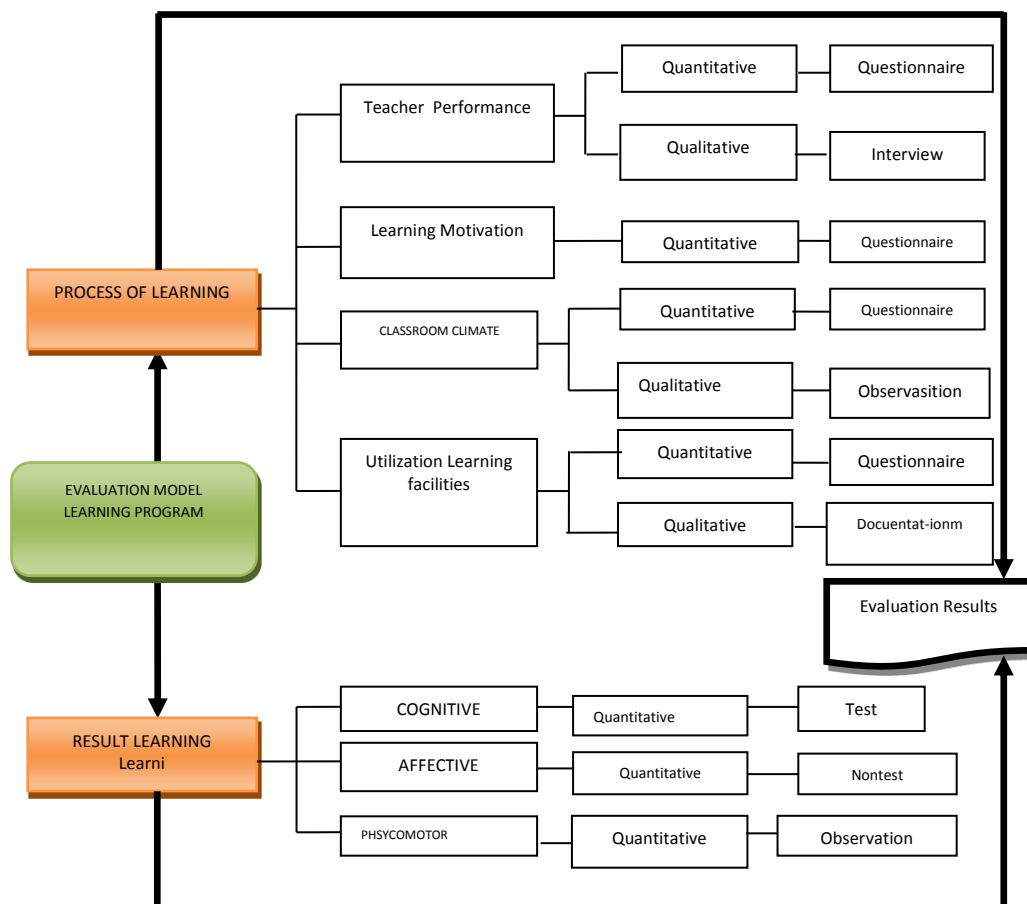
Figure 1. Model EPH

Based on the range of values Aiken's V between 0 and 1.00, then the Table 1, it could be said that the evaluation model program used has a good coefficient of content validity (content validity coefficient), because all Aiken's V above 0.7 (Azwar, 2013, p. 113).

The Trial Results EPH Model

*Test Limited*

Implementation of the limited trial conducted with 3 subjects trial namely headmaster and 13 teachers. The trial was designed to collect information related to the legibility of the model, in order to achieve the criteria of effective models, and practical. Information focused on the model regibility at the time used. The trial legibility evaluation models in the first phase were: (1) the legibility of the evaluation procedures, (2) the legibility of quality learning instruments, (3) instrument of learning outcomes, and (4) the legibility of the evaluation guidelines.

Assessment used a scale 5 the lowest score is 1, and the highest score 5. Based on the scores obtained on each item instrument, calculated the average score in points of the instrument. More results were presented in Table 2.

Table 2. Summary of Results Regibility Test Evaluation Model Program

| Number | Legibilty | Mean Score |
| --- | --- | --- |
| 1. | Evaluation Procedure | 4,55 |
| 2. | Process of Instrument | 4.45 |
| 3. | Results of Instrument | 4.40 |
| 4. | Evaluation Guidelines | 4.40 |

Table 2 showed all the components of program evaluation models in category score of more than 4, it means that all components of the program evaluation model included in good classification and could be used.

*Legibility Test Results Cognitive and Affective Domains*

Legibility tests about the cognitive and affective limited involving students class X and class XII, the number of students 15 people from three madrasas. Trial limited legibility was focused on aspects of the instructions clarity answering test, communicative language used, the choice of words that easy to understand, the structure of sentences was not complicated, and the sentences were not multiple interpretations. Assessment using a scale 5 a minimum score1, maximum score 5. The results of summery mean score legibility aspect of the instrument can be presented in Table 3.

Table 3. Summary of legibility the Cognitive and Affective Instrument test

| Number | Legibility | Mean Score |
|--------|-----------|------------|
| 1. | Cognitive domain test | 4,6 |
| 2. | Affective domain test | 4,6 |

The mean whole score legibility trial results, that there were legibility test instrument about the cognitive and affective category of the mean total score more than 4. The mean total of this score if converted to the assessment criteria of quantitative data into qualitative data with a scale of 5 was set into a good category test.

The Main trial

At primary trial, there are two instruments which have been tested, namely: (1) instrument of learning process that includes: teacher performance, learning motivation, classroom climate, and utilization of learning facilities, and (2) the evaluation instrument learning outcomes consist of:

instrument test questions cognitive, affective domain rubric tests and observation sheet rubric psychomotor.

Results of Instruments Learning Process

The main trial of the instruments learning process intended to indicate the feasibility of instruments, namely by estimating the instrument reliability and validity of the instruments used in this study. Based on the analysis with SPSS 19.00 for windows, can be indicated the index. The parameters used in testing validity by using the product moment correlation whereas reliability test with Cronbach Alpha. Validity viwed from coefficient r > 0.3 and reliability of Cronbach alpha coefficient > 0.7. Results summary SPPS all print out can be presented in table 4.

Table 4. Results of Testing Instrument Evaluation Model Learning Process

| Aspect | Loading Factor | *Alpha Cronbach* |
|--------|----------------|------------------|
| Teacher Performance | 0.551-0.642 | 0.731-0.793 |
| Learning Motivation (Students) | 0.430-0.637 | 0.719-0755 |
| Classroom Climate | Students 0.418-0.583 | Teachers 0.447-0.674 |
| Utility of learning Vasilities Fasilitas (Teacher) | 0.446-0.843 | 0.889 |

Table 4 shows all the instruments components of the model evaluation learning process are valid and reliable, because the loading factor > 0.3 and reliability of Cronbach alpha coefficient> 0.7.

Results Instruments Learning Outcomes

Trial learning achievement test consist of: instrument of cognitive domain, affective rubric, and psychomotor observation rubric sheet. Trial cognitive learning outcomes instrument, designed to indicate the feasibility model of the instrument, namely by estimating the instrument validity

and reliability of the instrument used in the study. The test results presented in Table 5.

Table 5.  Results of Testing Instrument Evaluation Model Learning Outcomes

| Instrument Test | loading factor | *Alpha Cronbach* |
|---|---|---|
| Cognitive | 0.413 - 0.806 | 0.750 - 0.789 |
| Affective | 0.418 - 0.680 | 0.714 - 0.788 |
| Psychomotor | - | 0.6196 - 0.7256 |

Table 5. The results of the loading factor and Cronbach Alpha showed all instrument components of learning outcomes is valid and reliable, because the coefficient r > 0.3 and reliability of Cronbach alpha coefficient > 0.7.

## The Results of Operations trials

CFA testing designed to evaluate the ability of the components were developed to be manifest, then used to reflect the variables evaluated. Two components tobe focus in learning program evaluation model of Islamic education is a learning process and results.

## Learning Process

The learning process was reflected by the three latent constructs, namely teacher performance, student motivation and classroom climate. Results of second order CFA

learning process viewed with some criteria Goodness of Fit is Root Mean Square Error of Approximation (RMSEA) <0:08; Chi-Square were obtained from the test has a probability greater than 0.05 (p> 0.05), and Normet Fit Index (NFI), Comparative Fit Index (CFI), and Googness of Fit Index (GFI)> 0.90. Visually model of CFA results were presented in Figure 2.
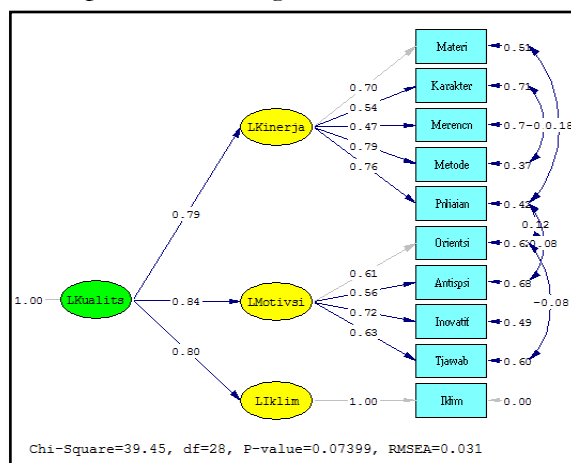


Figure 2.  CFA Second Order Learning Process

Manifests of a latent variable beside to reflect significantly also should be unidimensional. These properties were evaluated by testing consruct reliability. The test results consruct reliability (CR) learning process are presented in Table 15. The test results Consruct Reliability Scond Order can be presented in Table 6.

Table 6.  Test Results Consruct Reliability (CR) Second Order Learning Process

| No | Conctruct | Manifests | λ | E | CR |
|---|---|---|---|---|---|
| 1 | Teacher Performance | Materials | 0.700 | 0.5100 | |
| | | Character | 0.540 | 0.7084 | |
| | | Planning | 0.470 | 0.7791 | 0.7917 |
| | | Method | 0.790 | 0.3759 | |
| | | Rating | 0.760 | 0.4224 | |
| | | Orientation | 0.610 | 0.6279 | |
| 2 | Student learning motivation | Anticipation | 0.560 | 0.6864 | 0.7258 |
| | | Inovation | 0.720 | 0.4816 | |
| | | Responsibility | 0.630 | 0.6031 | |
| 3 | Classroom climate | climate | 1 | 0.0000 | 1 |

Based on these results known that the construct has a coefficient CR> 0.7, it means significant in reflecting latent constructs reflected, so that the components of the learning quality process in the learning program evaluation model of Islamic education was proper to use (Hair, Black, Babin, & Anderson, 2010, p. 92).

Learning outcomes

Learning outcomes reflected by the results of learning in cognitive, affective, and psychomotor domains. The implementation of learning program evaluation model be adapted with domain measured. Subject to the implementation of the model to evaluate the cognitive domain was not the same as the affective and psychomotor domains. Therefore CFA testing performed on each construct.

Cognitive Domain

Evaluation of cognitive lesson domain PAI in Madrasah Aliyah consisted of subjects Qur'an Hadits, Morals, Fiqh and Islamic Cultural History (ICH).

Knowledge of Al- Qur'an and the Hadis reflected by ten manifest grouped into three constructs, namely definition, basic understanding of the content, and functionality of the Qur'an. Second order CFA results are presented in Figure 3. The coefficient of chi-square = 41.31 with p value = 0.10198 > 0.05 indicated measurement model Qur'an-Hadith suitable with the population. When viewed from the other GOF parameters such as NFI, CFI and GFI more than 0.9. and RMSEA ≤ 0.08 indicates that the model was fit so it was not necessary to change the model. Visually model of CFA results are presented in Figure 3.

Akidah akhlak was reflected by fourteen manifest, which are grouped into four constructs, namely the principles of Akidah, Tauhid, Syiriq, and Morals. Second order CFA results were presented in Figure 4. The coefficient of chi-square = 83.78 with p value = 0.16167 > 0.05 indicate that measurement model of akidah akhlak suitable

with the population. GOF parameter values such as NFI, CFI and GFI more than 0.9. and RMSEA = 0.029 ≤ 0:08 indicates that the model fit. Visually model of CFA results are presented in Figure 4
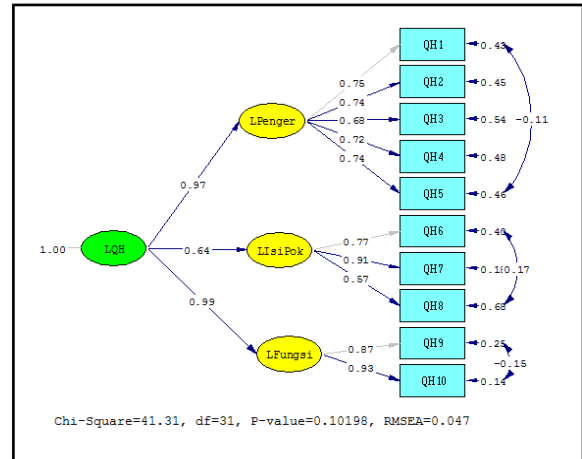


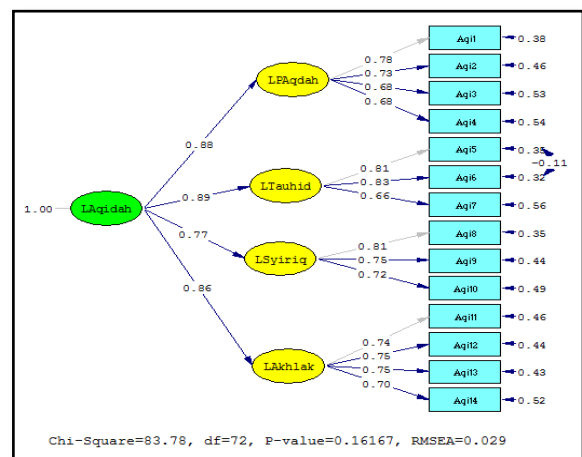Figure 3. CFA second Order Qur'an Hadith



Figure 4. CFA Second Order Akidah Akhlak

Fikih was reflected by twelve manifest grouped into four constructs, namely Principle of Worship, alms Law, sacriface Wisdom, and the pilgrimage law. Results of second order CFA was presented in Figure 5. The coefficient chi-square = 61.56 with p value = 0.17116 > 0.05 indicates measurement model of fikih was suitable with the population. GOF parameter values such as NFI, CFI and GFI more than 0.9. and RMSEA = 0.030 ≤ 0.08 indicated that the model fit. Visually model of CFA results are presented in Figure 5.

Islamic Cultural History was reflected by the ten manifest, which are grouped into

four constructs, namely: Understanding exemplary preaching Prophet in guiding the people, Understanding the leadership issue of Muslims after the Prophet's death, Understanding the probelem of islamic leadership after death Muhamad prophet, understanding the islamic development at the middle period clasical or the golden age and Understanding the development of Islamic in the medieval period/the dark age. Second order CFA results were presented in Figure 6. The coefficient of chi-square = 42.25 with p value = 0.06811 > 0:05 indicates that measurement model of Islamic Cultural History fits with the population. GOF parameter values such as NFI, CFI and GFI more than 0.9. and RMSEA = 0.064 ≤ 0:08 indicates that the model fit. Visually model of CFA results are presented in Figure 6.
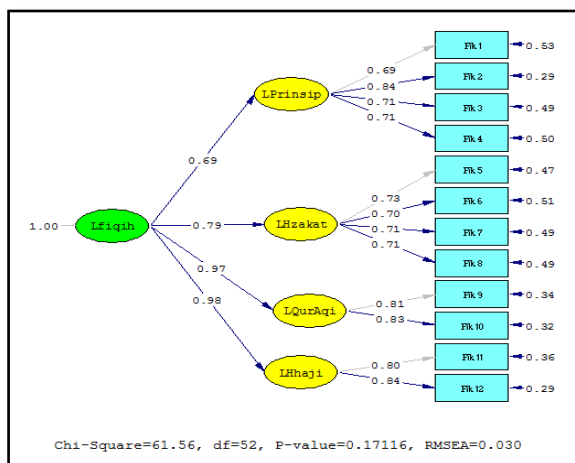


Chi-Square=61.56, df=52, P-value=0.17116, RMSEA=0.030

Figure 5. CFA Second Order Fikih
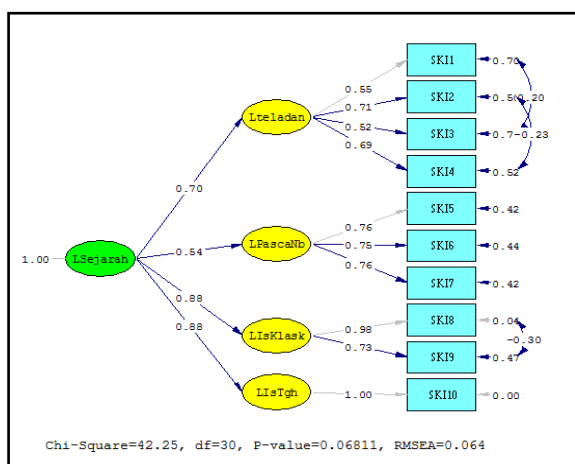


Chi-Square=42.25, df=30, P-value=0.06811, RMSEA=0.064

Figure 6. The Second Order CFA SKI

Testing unidimensionality at the second order gained coefficient Consruct reliability > 0.7 for each construct. it Means that each manifest used to reflect the construct was unidimensional (Hair et al., 2010, p. 92). Results Consruct Reliability Second Order. Learning Outcomes Cognitive Domains were presented in Table 7.

Table 7. The Results of Consruct Reliability Test (CR) Scond Order Learning Outcomes Cognitive Domains

| Construct | λ | E | CR |
|---|---|---|---|
| Al-Qur'an Hadis | 0.57-0.93 | 0.1351-0.6751 | 0.8050-0.8250 |
| Akidah Akhlak | 0.66-0.83 | 0.439-0.5376 | 0.8050-0.8250 |
| Fikih | 0.69-0.84 | 0.2944-0.5239 | 0.84-0.828 |
| SKI | 0.52-1 | 0-0.7296 | 1- 0.714 |

Based on the significant results of the manifest and latent constructs reflected; unidimensional nature manifest in the second group order, and constructs in the first order. Then the Islamic Cultural History component in the learning program evaluation model of Islamic education is proper to be used.

Affective Domain

Affective domain was reflected by twelve manifest grouped into four constructs, namely: Interests (label: L intrestst), Attitude (label: L attitude), Discipline (label: LDicipline), and Cooperation (label: L cooperation). Results of second order CFA wa presented in Figure 7. The coefficient of chi-square = 118.66 with p value = 0.08680 > 0:05 indicated that measurement model of affective domain was suitable with the population. GOF parameter values such as NFI, CFI and GFI more than 0.9. and RMSEA = 0.023 ≤ 0:08 indicates that the model fit. Visually model of CFA results were presented in Figure 7.

Testing unidimensionality at the second order gained coefficient Consruct reliability > 0.7 for each construct. it Means that each manifest used to reflect the construct

was unidimensional (Hair et al., 2010, p. 92). Results Consruct Second Order Reliability Domains Affective presented in Table 8.
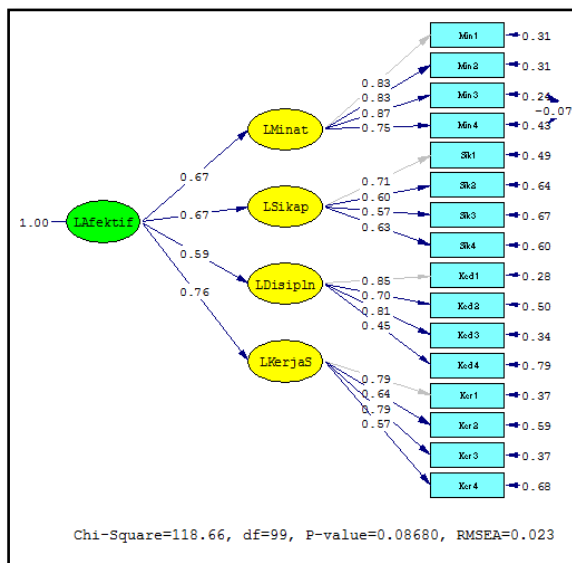


Figure 7. CFA Second Order Affective

Table 8. The results of Consruct Reliability Tes (CR) Scond Order Affective Domains

| No | Construct | Manifest | λ | E | CR |
|----|-----------|----------|------|--------|-------|
| 1 | Linterest | Min1 | 0.83 | 0.311 | 0.892 |
| | | Min2 | 0.83 | 0.311 | |
| | | Min3 | 0.87 | 0.2431 | |
| | | Min4 | 0.75 | 0.438 | |
| 2 | Lattitude | Sik1 | 0.71 | 0.496 | 0.723 |
| | | Sik2 | 0.6 | 0.640 | |
| | | Sik3 | 0.57 | 0.675 | |
| | | Sik4 | 0.63 | 0.603 | |
| 3 | Ldicipline | Ked1 | 0.85 | 0.278 | 0.804 |
| | | Ked2 | 0.7 | 0.510 | |
| | | Ked3 | 0.81 | 0.344 | |
| | | Ked4 | 0.45 | 0.798 | |
| 4 | Lcooperating | Ker1 | 0.79 | 0.376 | 0.794 |
| | | Ker2 | 0.64 | 0.590 | |
| | | Ker3 | 0.79 | 0.376 | |
| | | Ker4 | 0.57 | 0.675 | |

Based on the significant results of the manifest and latent variables constructs was reflected, unidimensional nature manifest in the second group order. Then the affective component in the learning program evaluation model of Islamic education is proper to be used.

Psychomotor Domain

Psychomotor learning outcomes was presented through the practice of ablution and prayer. The measurement is reflected by nineteen manifest that grouped in two constructs, namely: the practice of ablution and prayer. the Second order CFA results were presented in Figure 8. The coefficient of chi-square = 148.08 with p value = 0.09726 > 0.05 indicates that Psychomotor Domains measurement model was suitable with the population. GOF parameter values such as NFI, CFI and GFI more than 0.9. and RMSEA = 0.041 ≤ 0.08 indicated that the model fit. Visually model of CFA results were presented in Figure 8.
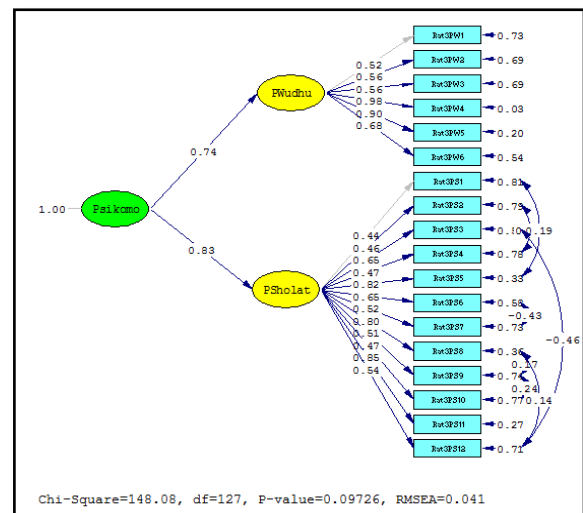


Figure 8. CFA Second Order Psychomotor

Testing unidimensionality at the second order gained coefficient Consruct reliability> 0.7 for each construct. Means each manifest used to reflect the construct is unidimensional (Hair et al., 2010, p. 92), as the result of second order reliability construct ablution and prayer practices were presented in Table 9.

Based on the significant results of the manifest and latent constructs was reflected unidimensional nature of the group's manifest in the second order, and constructs the first order. So the psychomotor component on evaluation model of Islamic education learning programs is proper to be used.

Table 9. The Results of Construct Reliability Test (CR) Sconde Order Learning Process

| Number | Construct | Manifest | $\lambda$ | E | CR |
|---|---|---|---|---|---|
| 1 | Ablution Practice | Rat3PW1 | 0.52 | 0.7296 | |
| | | Rat3PW2 | 0.56 | 0.6864 | |
| | | Rat3PW3 | 0.56 | 0.6864 | |
| | | Rat3PW4 | 0.98 | 0.0396 | 0.860 |
| | | Rat3PW5 | 0.90 | 0.19 | |
| | | Rat3PW6 | 0.68 | 0.5376 | |
| 2 | Praying Paractice | Rat3PS1 | 0.44 | 0.8064 | |
| | | Rat3PS2 | 0.46 | 0.7884 | |
| | | Rat3PS3 | 0.65 | 0.5775 | |
| | | Rat3PS4 | 0.47 | 0.7791 | |
| | | Rat3PS5 | 0.82 | 0.3276 | |
| | | Rat3PS6 | 0.65 | 0.5775 | |
| | | Rat3PS7 | 0.52 | 0.7296 | 0.874 |
| | | Rat3PS8 | 0.80 | 0.36 | |
| | | Rat3PS9 | 0.51 | 0.7399 | |
| | | Rat3PS10 | 0.47 | 0.7791 | |
| | | Rat3PS11 | 0.85 | 0.2775 | |
| | | Rat3PS12 | 0.54 | 0.7084 | |

Study of the Last Product

Model of EPH has been tested, both quantitative and qualitative. The trial results quantitatively in medium scale by using SPSS program and field operations (operational field testing) used CFA confirmatory analysis program, the results showed that the model evaluation instrument EPH it have been fullfilled reliability coefficient, the items of have a instrument were valid, and the model was fit. Further test results are qualitatively EPH model evaluation, the results showed that the implementation of the evaluation model was practical, effective, and efficient.

Evaluation EPH models have been equipped components and evaluation procedures, the learning process instruments, instrument results (outputs) of learning, evaluation guidelines was very detailed and clear. Thus the model was applied in Madrasah Aliyah (MA), MTs (MTs), can also be applied to public schools (SMA and SMP).

Characteristics of Evaluation Model

(1) The model is used to evaluate the learning Islamic education program in Madrasah Aliyah

(2) The Evaluation model is the internal evaluation conducted by headmater and teachers to monitor the learning process and learning outcomes

(3) the Model evaluation learning program has two components evaluating learning programs, namely the evaluation of the learning process and evaluation of learning outcomes.

(4) Using of this model is not dependen on a particular learning program evaluation model that implemented by headmaster and PAI teachers.

(5) EPH model can be applied to Curriculum 2013

(6) This model is open for further development, and have not become the final product.

Implication of Evaluation Model

(1) Evaluation EPH models is very complete and comprehensive to evaluate learning process that includes evaluating teacher performance, student motivation, classroom climate, and utilization of learning facilities, conducted headmaster and teachers PAI

(2) Evaluation EPH models is very complete and comprehensive to evaluate learning outcomes include: evaluation of learning outcomes cognitive, affective, and psychomotor, committed teachers Islamic education

(3) This model is very practical to use, easy tobe understood by the user, and easy tobe implemented at the islamic school or public school.

Research Limitations

(1) The evaluation process have not involved the external parties that are independent but rely solely on the ratings of internal madrassas so it is possible there is an element of subjectivity assessment of implementing.

(2) Requires substantial costs mainly related to the implementation of testing, validation experts, practitioners, as well as regibility test. High financial also influenced by the location of the research is so far.

(3) Limitations of validator involved in this study were derived from a limited number of PAI teachers.

**Conclusions**

Conclusion based on results of this research consists of: (1) evaluation model of Islamic Education Learning program developed consist of: (a) the evaluation procedures, (b) the evaluation instrument (c) the evaluation guide; (2) according to experts, PAI Teachers and headmaster of Madrasa, procedures, instruments, and evaluation guidelines are already well developed, 3) The instrument developed entirely valid (load factor> 0.3), reliable (CR> 0.7), and qualify as a fit model (RMSEA ≤0.08 and NFI, CFI and GFI> .90). and 4) the assessment of the experts, PAI teacher and headmaster, EPH models asserted effective, practical, and easy to use, supported with a valid and reliable instrument.

Based on the conclusions of research, socould be suggested that: (1) the user should read and understand guidelines to the models that have been prepared, (2) the instruments are given to students adjusted to the phase of the evaluation is continuing, time, and the right conditions, (3) using models EPH, can be adapted to the purpose, useful and level of desired program, (4) in learning process teachers should be able to be a good guidance, maintain the performance, creating a conductive classroom climate, motivating, utilize of learning facilities in order to reach the good learning students achievements, and (5) this product should be used as guidelines in evaluating the learning process and learning outcome PAI because it able to reach the side of the cognitive, affective, and students' psychomotor comprehensively.

**References**

Agustian, A.G. (2003). *ESQ power*. Jakarta: Arga.

Azwar, S. (2013). *Reliabilitas dan validitas* (4th ed.). Yogyakarta: Pustaka Pelajar.

Borg, W. R., Gall, J. P., & Gall. M. D. (1993). *Applying educational research: A practical guide*. New York: Longman.

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis* (7th ed.). New York: Pearson Prentice Hall.

Mardapi, D. (2000). Evaluasi pendidikan. In *Konvensi Pendidikan Nasional 19-23 September 2000*. Universitas Negeri Jakarta.

Morrison, D. M., Mokashi, K., & Cotter, K. (2006). *Instructional quality indicator: Research foundations*. Cambrigde: Cambrigde University. Retrieved Mei 13, 2011, from www.co.nect.net.

Muhaimin. (2009). *Rekonstruksi pendidikan islam (dari paradigma pengembangan, manajemen kelembagaan, kurikulum hingga strategi pengembangan)*. Jakarta: PT Grasindo Persada.

Norušis, M. J. (1986). *SPSS/PC+ for theimbbc/xc/at*. Chicago: SPSS Inc.

Sudjana, N. (2002). *Dasar-dasar proses belajar dan mengajar*. Bandung: Sinar Baru Algesindo.

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

Sutrisno. (2006). *Pendidikan islam yang menghidupkan*. Yogyakarta: Kota Kembang.

Wahyudi. (2003). Penyusunan dan Validasi Kuesioner Iklim Lingkungan Pembelajaran di Kelas. *Jurnal Pendidikan dan Kebudayaan, 043*.

Zamroni. (2005). Mengembangkan kultur sekolah menuju pendidikan yang bermutu. In *Seminar Nasional Mengembangkan Kultur Sekolah di Yogyakarta* Retrieved November 23, 2005.