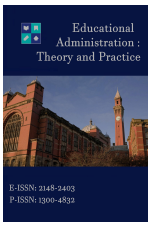
Kuram ve Uygulamada Eğitim Yönetimi
Educational Administration: Theory and Practice
2023, Cilt 29, Sayı 2, ss: 300-313
2023, Volume 29, Issue 2, pp: 300-313
www.kuey.net



Development of Evaluation Model of Character-based Islamic Education Program in Elementary Schools

Siti Choiriyah [1*], *Khuriyah* [2], *Hakiman* [3]

Article History

Article Submission
11 November 2022
Revised Submission
08 December 2022
Article Accepted
14 February 2023

Abstract

The success of the Islamic Education program and character development through the program can be determined through evaluations. The objective of this study is to develop an evaluation model of character-based Islamic Education learning programs. This model will provide more accurate information to Islamic Education teachers and school leaders, enabling the improvement of character-based Islamic Education programs. This research applied Research & Development model incorporating the Plomp development model and the Borg & Gall model. The evaluation model developed, the character-based education in Islamic Education (EPK-PAI), utilized the process input and output evaluation approach. This study was conducted in all Elementary Schools in Kartasura Sukoharjo sub-district. The data collection technique includes a questionnaire, FGD, and the Delphi technique. Research instruments were analyzed with Confirmatory Factor Analysis and the EPK-PAI model was analyzed descriptively based on the assessment of Islamic Education teachers and school leaders. The EPK-PAI model was proven using Structural Equation Modeling. This research has developed an evaluation model of character-based Islamic Education programs in Elementary Schools using systematic steps from the Borg & Gall model, including information gathering, planning, initial product development, initial field testing, product revision, and main field testing. The EPK-PAI model in this study comprises four categories of evaluation that are classified into three dimensions, namely input, process, and output. The EPK-PAI model has reliable construct validity and reliability, thereby rendering it suitable for evaluating character-based Islamic education programs in the elementary school context.

Keywords: Evaluation Model Development; Islamic Education Learning; Character Education

[1*]*Associate Professor, Faculty of Education, UIN Raden Mas Said Surakarta, Surakarta, Indonesia, siti.choiriyah@staff.uinsaid.ac.id*

[2]*Associate Professor, Faculty of Education, UIN Raden Mas Said Surakarta, Surakarta, Indonesia, khuriyah@staff.uinsaid.ac.id*

[3]*Assistant Professor, Faculty of Education, UIN Raden Mas Said Surakarta, Surakarta, Indonesia, hakiman@staff.uinsaid.ac.id*

Introduction

Character education is the foundation for achieving the goals of national education to create a society with noble, moral, ethical, cultured, and civilized characters based on the philosophy of Pancasila. Character education is an effort to support the realization of the ideals outlined in Pancasila and the Preamble of the 1945 Constitution. Recent issues have highlighted the importance of character education as a foundation for educational development. The National Long-Term Development Plan (RPJPN 2005-2015) reveals the implicit ethos of the government by placing a salient emphasis on the cultivation of personal character as a paramount initiative for national advancement. Adopting the concept of Thomas Lickona (2004), the character is the harmony of moral knowing, moral feeling, and moral action. This implies that an individual possessing integrity exhibits cognitive alignment with moral principles, emotional alignment with ethical values, and behavioral alignment with virtuous actions.

Islamic Education (PAI) is a conscious and planned effort in preparing students to recognize, understand, appreciate, have faith, act virtuously, possess a noble character, and practice the Islamic teachings based on the Al-Qur'an and al-Hadith through various instructional methods, training, mentoring and through the incorporation of practical experience (Ramayulis, 2005). Islamic Education learning is a very prerequisite program with character values to help students increase their understanding, appreciation, and practice of Islamic teachings so that they become more devoted to Allah and have a strong moral character in their personal, community, national, and state life (Minister of Religious Affairs Regulation Number 2 of 2008).

The success of the Islamic Education program and character development through the program can be determined through evaluations. The evaluations provide information that can be used to make decisions, create policies, and plan future programs. To make the most of the information, evaluations must be thorough, accurate and reliable, and provided promptly.

Educational evaluation, when viewed in terms of targets, is divided into macro and micro evaluation. The primary objective of the macro evaluation is to enhance premeditated endeavors, while micro evaluations serve to assess pedagogy within the classroom setting (Mardapi, 2000). Teachers are responsible for creating and implementing classroom learning programs, while school leaders are responsible for evaluating programs that have been created and implemented by the teacher.

According to a survey conducted by Widoyoko (2007) in several schools of Central Java, the participants which include teachers, principals, and vice-principals reported that the effectiveness of existing educational activities is primarily determined by the achievement of learning outcomes. Meanwhile, the evaluation of the quality of the learning process has received less attention. Assessing learning outcomes tends to be focused more on academic or cognitive abilities, while less attention is paid to emotional or character-building skills, as well as psychomotor abilities. Although the survey results may not provide a comprehensive representation of all learning outcomes assessments in Indonesia, it highlights that there are deficiencies in evaluating learning programs, specifically concerning character-based Islamic Education programs that have been running so far.

The research suggests the development of an evaluation model for character-based Islamic Education programs in schools. The evaluation model of the character-based Islamic Education program can provide more detailed information for teachers and school leaders, and it can be used to improve the character-based Islamic Education program.

Prior investigations on the evaluation models of character education programs in Indonesia, specifically in Surakarta, have utilized the Stuflebeam survey model within the context of Junior high schools (Nurindarwati, 2022). Other studies evaluated character education programs in Islamic schools (Anidi, 2017; Jumadi et al., 2022; Muhlisin et al., 2019). Differently, this current study incorporated the Plomp development model and the Borg & Gall model. This study, therefore, fills a gap in the literature by providing a novel examination of an evaluation model of character-based Islamic Education programs within the context of elementary schools using the combination of the Plomp development model and the Borg & Gall model, which has hitherto been unexplored by previous researchers. The objective of this study is to develop an evaluation

model of character-based Islamic Education learning programs.

Literature Review

Development of Evaluation Model

The understanding of the definition of program evaluation may differ based on the various definitions provided by evaluation experts. For example, Stufflebeam, as cited by Ansyar (1989), defines evaluation as a process of acquiring and presenting appropriate information for decision-making. The Joint Committee on Standards for Educational Evaluation (1994) defines it as a systematic investigation of the success of a goal, while Djaali et al. (2000) describe it as a process of judging something based on established objective standards to guide decisions about the object being evaluated.

Rutman and Mowbray, 1983 propose that evaluation is the use of scientific methods to assess the implementation and outcomes, which is beneficial for the decision-making process. Chelimsky (1989) defines evaluation as a systematic research method for assessing the design, implementation, and effectiveness of a program. Wirawan, 2006 asserts that evaluation is the process of gathering and presenting information on various evaluation objects. It evaluates the object based on evaluation standards and the results are used to make decisions about the evaluation objects. From the evaluation definition above, it can be inferred that evaluation is the application of a systematic scientific procedure to conduct a design assessment. It provides information for making decisions about the implementation and effectiveness of a program.

Program evaluation is the first step in supervision. It allows to a collection of accurate data, which can then be used to provide targeted guidance. Program evaluation is critical and beneficial, particularly for decision-makers. This is because the findings from the program evaluation can be used by decision-makers to determine the next steps for the program currently being implemented or that has already been carried out.

Development research, as outlined by Borg and Gall, 1983, is the process to develop and validate educational products. It is further expounded by Borg and Gall that the steps that must be taken in development research include: 1) Research and information collection, 2) Planning, 3) Development of a preliminary product form, 4) Initial field testing, 5) Main product revision, 6) Main field testing, 7) Final product revision, 8) Final field testing, 9) Dissemination & Implementation. Plomp, 1997 also states that development research comprises the following stages: initial examination, design, realization/construction, testing, evaluation, and revision and implementation.

To research the development of an evaluation model, it is imperative to ascertain the types of program evaluation models utilized by experts:

CIPP Model (Context-input-process-product).

The CIPP model was developed by Stufflebeam in 1971 (Cates, 1990). It is an acronym for the "Context, Input, Process, and Product" evaluation model. The CIPP model (1971) views evaluation in four dimensions: the context dimension, the input dimension, the process dimension, and the product dimension. It is a systematic process for evaluating educational programs, policies, and projects. The uniqueness of this model is that it relates each type of evaluation to a decision-making tool for the planning and operation of a program. The CIPP model offers a fairly comprehensive format for each stage of evaluation, beginning with context, input, process, and product. The CIPP model is widely used in the educational evaluation as it covers all aspects of the program, from its planning to implementation and its outcomes. The CIPP model is composed of four phases. Context evaluation: This phase entails assessing the program's setting and context, as well as the needs and characteristics of the target population. Input evaluation: This phase involves assessing the program's resources, materials, and personnel. Process evaluation: During this phase, the program's implementation is evaluated, including how well it is delivered and how well participants engage with it. Product evaluation entails assessing the program's outcomes, including whether or not it met its goals and objectives.

Kirkpatrick Model

Kirkpatrick's model comprises four steps: reaction, learning, behavior, and results. This model has several benefits when compared to other models. Firstly, it is considered more comprehensive as it includes cognitive, skill, and affective aspects, and it not only evaluates learning outcomes but also includes processes, outputs, and outcomes. Secondly, it is suitable for the classroom level, as it does not involve a lot of other parties in the evaluation process (Kirkpatrick, 1998).

Islamic Education Learning

Definition of Islamic Education Learning

Islamic education is a conscious and planned effort in preparing students to recognize, understand, appreciate, belief in, have reverence, have a noble character, and practice Islamic teachings of the main source of the Al-Qur'an and Al-Hadith, through guidance, teaching activities as well as practical experiences (Ramayulis, 2005).

According to Azizy as cited in Majid & Andayani (2004), the essence of education is the transfer of values, knowledge, and skills from the older generation to the younger generation. Islamic Education has two objectives, which are to educate students to have good morals and to acquire knowledge and understanding of Islam.

In summary, Islamic Education can be understood as a process of developing students' potential through the teachings of Prophet Muhammad SAW, based on the Al-Qur'an and Al-Hadith, to become human beings who fear Allah.

Scope of Islamic Education

The scope of Islamic education includes harmony, conformity, and balance between human relationships with Allah SWT, oneself, other humans, and their environment.

Therefore, it can be inferred that the scope of Islamic Education at all levels of education (from elementary school to higher education) is integral to the five units aforementioned, namely Al-Quran, Aqidah, Sharia, Morals, and Tarikh (Ramayulis, 2006). In other words, the Islamic education program should cover Al-Quran, Aqidah, Sharia, Morals, and Tarikh. Together, these five units are expected to create future generations with a good relationship with Allah, fellow humans, and all creations.

Character Education

Definition of Character Education

According to Binks et al. (2016), the human character is an embodiment of one's habits. Meanwhile, Meyer (2020) defines character as a quality that is consistently demonstrated by oneself. Character is shaped by both heredity and the environment. Both of these factors play a role in shaping students' characters, thus, an individual's habits will shape their character.

Meanwhile, Musfiroh (2008) defines character as a series of attitudes, behaviors, motivations, and skills. The character comes from Greek which means to mark and to focus on applying the value of goodness in the form of action or behavior.

In a special context, character is also defined as a person's character, behavior, morals, or personality formed from the results of internalization as virtues believed and used as a basis for point of view, thinking, behaving, and acting (Pusbuk, 2011).

Human life is full of virtue. According to Lickona (1992), human life is divided into two categories; self-oriented virtuous such as self-control and moderation, and other-oriented virtues, such as a willingness to share (generosity) and feel goodness (compassion).

According to Ramli, 2003 character education in substance and meaning is similar to moral and attitude education aiming to shape children's personalities and become good human beings, communities, and citizens. The standards for what constitutes a good person, society, and citizen are often shaped by the culture of the community and its nation.

Value of Character Education

The 18 character values are formulated based on the values of Pancasila, the 1945 Constitution, Indonesian culture, and the National Education System. They are religious, honest, tolerance, disciplined, hard work, creative, independent, democratic, curiosity, national spirit, love

of the homeland, achievements appreciation, friendly/communicative, love of peace, love of reading, environmental care, social care, and responsibility (Ministry of National Education, 2010).

The teacher can augment or reduce these values according to the needs of the community served by the school and the nature of the basic competency/competency standards and the subject matter of a learning subject. However, according to Budimansyah (2011), six values such as honesty and responsibility, intelligence, creativity, perseverance, and care are expected to be the minimum values developed in every school. Meanwhile, in the 2013 curriculum, character values are further developed by achieving core competencies of spiritual attitudes and social attitudes.

Methodology

Research Method

Based on the primary paradigm of humanities and social science research, quantitative research, qualitative research, and mixed research methods are studied. However, there are also four studies without clear research methods, which are generally non-empirical studies, usually proposing, expanding, or applying strategies and theories.

Research Topic Coding

This research employs Research and Development model utilizing a development model by Plomp combined with Borg & Gall model. The definition of a combination in this case refers to the research and development steps, which primarily draw upon Plomp's model. Additionally, the determination of the number of test subjects adheres to the guidelines provided by Borg & Gall, increasing the number of test subjects from the first stage to the next.

Meanwhile, the evaluation model developed used an input, process, and outcome approach. Evaluation of the learning process in this case is called evaluation or learning quality. This model is focused on assessing the output of learning and is called the Evaluation Model of Character-based Islamic Education Learning (EPK-PAI model). This model is based on the assumption that the outcomes of Islamic Education learning can be divided into three outputs, including cognitive, affective/character, and skill.

The EPK-PAI model is a combination of Stufflebeam's CIPP (Context, Input, Process, Product) model and Kirkpatrick's evaluation model, with some modifications and additions to certain aspects of evaluation. The components of the development of the evaluation model for character-based Islamic Education in Primary Schools with EPK-PAI include learning inputs such as teacher competencies and learning tools, learning processes such as teaching methods and classroom management, and learning outputs that include cognitive, skills, and psychomotor.

The development of the evaluation model of the character-based Islamic Education learning program in this school took the following steps: (1) analysis of theory and results of previous research, (2) design preparation based on the analysis of evaluation models from previous research, (3) expert judgment, which was carried out through focus group discussions (FGD) and the Delphi technique, (4) product testing, and (5) data analysis. Experts involved in FGD activities consist of evaluation experts, character education experts, Islamic Education learning experts, Primary School Islamic Education teachers, and Primary School principal/Primary School deputy prinExperts participating in focus group discussions include evaluation experts, character education experts, Islamic Education experts, primary school Islamic Education teachers, and primary school principals/assistant principals.

Product Trials

Aspects validated in the evaluation model test include a) Evaluation of the EPK-PAI model, and b) Evaluation of instruments of the EPK-PAI model. The study participants were primary school students, Islamic Education teachers, and the school principal as a representative of school leaders.

The validity test in this study refers to empirical validity. The study used a statistical method

called factor analysis, specifically the SPSS version 23 program's exploratory factor analysis (EFA) to test the construct validity of the teacher competencies, learning infrastructure, and student learning outcomes.

The Technique of Collecting Data

The study was conducted by using a combination of survey questionnaires, focus group discussion, and the Delphi technique, which is a method of gathering input from selected experts to reach a consensus after the focus group discussion (Linston & Turrof, 2002).

Data Analysis Measure

The data collection tools were analyzed using Confirmatory Factor Analysis (CFA). The EPK-PAI model was analyzed descriptively by looking at the assessment of Islamic Education teachers and school leaders, using Structural Equation Modeling (SEM) to test the assumptions of the EPK-PAI model. Both the measurement model and evaluation model were analyzed using the Lisrel 8.30 program. Four indicators were used to test the fit between the theoretical model and the empirical data: (1) Chi-Square, (2) Significance Probability, (3) Root Mean Square Error of Approximation (RMSEA), and (4) Goodness of Fit Index (GFI). The quantitative data obtained from the assessment instrument was analyzed to find the mean score, which was then converted to a qualitative scale of 5, and described qualitatively (Sudijono, 2003).

Results

The development of the evaluation model of the character-based Islamic education program was carried out by adopting and modifying the development model from Borg and Gall through the following steps:

Research and Information Collection

In this step, the researchers began by reviewing the literature on the subject, studying relevant research findings, and conducting initial research. The literature review and examination of relevant research results involved looking at the CIPP model from various sources such as books, journals, articles, and research findings, to identify and understand evaluation methods that could be applied to evaluating character-based Islamic Education programs. After that, the researchers looked into the standards for incorporating character education into the KTSP and K-13 curriculum in schools.

Conduct an analysis,

The next step in gathering information was to conduct an analysis, which looked at the evaluation model that needs to be developed. The participants in this study were the principal, Islamic Education teachers, regular classroom teachers, and parents of elementary school students. They were selected from state primary schools in Kartasura. The data was collected through interviews. The data obtained were analyzed descriptively. The results of the preliminary research indicate that the participants demand an evaluation model for the character-based Islamic Education learning program.

Planning (Preparing the Draft Design of the Initial Model)

The activities in the planning stage are:

To create a prototype evaluation model for the character-based Islamic Education learning program, the result of this stage is a prototype evaluation model. The prototype is then organized into a grid format as an evaluation model instrument. The stages or procedures for evaluating the character-based Islamic Education learning program include the evaluation of inputs, processes, and outputs. Evaluation of the input quality consists of evaluating the teachers' competence and infrastructure. Evaluation of the process quality consisted of evaluating the learning process and class management, while evaluation of the output quality was done by evaluating the competence of graduates.

To prepare a prototype of the instrument assessing the developed evaluation model, the instrument includes two aspects, namely: (1) the design or layout of the evaluation model being developed, and (2) the substance of the evaluation model.

Developing a Preliminary Form of Product

Following the completion of the planning stage, the next stage is to build a model application that can be tested in the field. At this stage, the model design is completed by confirming with Islamic Education experts, practitioners, and Evaluation experts whether or not important indicators should appear in the evaluation model developed using FGD and Delphi techniques. Technically, the Delphi technique was carried out in two rounds by meeting the respondents directly. The Islamic education practitioners who contributed were Ms. Sulasmi, S.Pd (Principal of Pucangan Elementary School), and Mr. Nugroho, S.Pd. (Religion Teacher at State Elementary School/SDN 4 of Kartasura). Based on the draft of the instrument grid, the overall draft of the instrument grid is accepted by practitioners.

Main Field Testing

Main field testing was conducted in eight Primary Schools in Kartasura, namely State Elementary School (SDN) of Pucangan 1, SDN Pucangan 5, SDN Kartasura 1, SDN Kartasura 4, SDN Kartasura 6, SDN Ngadirejo 1, SDN Ngadirejo 4, dan Integrated Islamic Elementary School (SDIT) of Al-Kautsar.

Validity Test in EPK-PAI Model

Construct Validity of Teacher Competency

At this stage, the validity test is empiric validity using factor analysis, specifically the SPSS 23 program exploratory factor analysis (EFA). The feasibility of using factor analysis is the first requirement of factor analysis. The first requirement is that the number of observations (data) is sufficient for factor analysis if the Kaiser Meyer Olkin (KMO) value is greater than 0.5. Table 1 shows the results of KMO and Barlett's Test.

Table 1. KMO and Barlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy		.790
Bartlett's Test of Sphericity	Approx. Chi-Square	102.939
	Df	6
	Sig.	.0001

The KMO value is (0.790), which is greater than 0.5, indicating that the first requirement for data adequacy has been met. Furthermore, the second condition demonstrates that this factor analysis is feasible when the multivariate correlation with Bartlett, Sig. is tested. It must be less than the standard Alpha of 0.05. It can be seen from the output that Sig 0.00 is lower than Alpha 0.05 making the null hypothesis (rho = 0) rejected and the alternative hypothesis (rho ≠ 0) accepted. Thus, it can be concluded that there is a correlation among multivariate variables (Table 2).

Table 2. Anti-image Matrices

		personality	**pedagogy**	**professional**	**social**
Anti-image Covariance	personality	.621	-.124	.042	-.026
	pedagogy	-.124	-.166	-.092	-.046
	professional	.042	-.092	.153	-.094
	social	-.026	-.046	-.094	.211
Anti-image Correlation	personality	.834[a]	-.387	.137	-.071
	pedagogy	-.387	.755[a]	-.575	-.244
	professional	.137	-.575	.743[a]	-.521
	social	-.071	-.244	-.521	.838[a]

Furthermore, the correlation value among multivariate variables can be seen in Table 2 of the output on Anti Image Matrices. We need to pay attention to the Measure of Sampling Adequacy (MSA) only, which is the letter "a" in the Anti Image Correlation.

It was found that all MSA is greater than 0.5. Thus, it can be analyzed further. The next output is shown in Table 3.

Table 3. Communalities

	Initial	**Extraction**
personality	1.000	.499
pedagogy	1.000	.907
professional	1.000	.877
social	1.000	.858

This Communality can later be used as a validity test which can explain the existing variables. Based on the output, there are no new factors formed in the variable of teacher competency. The formed factors explain that the variable of personality competency is (49.9%), the variable of pedagogical competency is (90.7%), and so on until the last variable is the variable of social competence. The next output is Total Variance Explained by Factor(s) as shown in Table 4.

Table 4. Total Variance Explained

Component	**Initial Eigenvalues**			**Extraction Sums Squared Loadings**		
	Total	**% of Variance**	**Cumulative %**	**Total**	**% of Variance**	**Cumulative %**
1	3.141	78.525	78.525	3.141	78.525	78.5252
2	.610	15.256	93.782			
3	.152	3.807	97.589			
4	.096	2.411	100.000			

Extraction Method: Principal Component Analysis

Many factors may be formed from the variables. Factor I explains the data variation of (78.525%). This value is obtained from 3.141 per number of variables multiplied by 100% so that 3.141 / 4 * 100% = 78.525%. Then factor II, in the same way, explains the variation of the data by 15.256%. With the Eigenvalue, the set value is higher than 1. If the standard Eigenvalue is > 1, the result is only one factor formed (unidimensional) as in Table 5.

Table 5. Component Matrix[a]

	Component
	1
Personality	.706
Pedagogy	.952
Professional	.936
Social	.926

Extraction Method: Principal Component Analysis.

a.1 component extracted

The Component Matrix shows how much correlation exists between each variable and the formed factors (only one). It reveals that they are all strongly correlated with one formation factor (component); all are greater than 0.5. Because the resulting factor is only one, it cannot be rotated to find the best correlation, allowing a variable to be determined to which factor it belongs.

Because it employs extraction with an Eigenvalue greater than one, only one factor is formed. All variables have a strong correlation with this one formation factor, according to the Component Matrix (Unidimension). So, it is concluded that there is only one factor (Teacher Competence) reflecting the four variables (personality competency, pedagogy competency, professional competency, and social competency). This means that the instrument on teacher competency is unidimensional and empirically valid. As a result, it can be used to create a model for evaluating character-based Islamic education in primary schools.

Construct Validity of Learning Infrastructure

KMO and Barlett's Test is derived from the empirical test. It can be seen from the output that

Sig 0.00 is lower than Alpha 0.05, so the null hypothesis (rho = 0) is rejected and the alternative hypothesis (rho ≠ 0) is accepted. It can be concluded that there is a correlation among multivariate variables. Based on the anti-image table, all MSA is equal to 0.5, so it can be investigated further. The next output is commonalities. Based on the output, there are no new factors formed in the variables of infrastructure and facilities. The variables explain that the public infrastructure variable is 78.1%, and the classroom variable is 78.1%. The total Variance Explained by Factor is the next output (s). It can be seen that the formed Factor I can explain the data variation by 78.07%. It is obtained from 1.561 per number of variables multiplied by 100% so that 1.561 / 4 * 100% = 78.070%. Then factor II, in the same way, explains the variation of the data by 21.930%.

When the Eigenvalue is used, a value greater than one is determined. If the standard Eigenvalue is greater than one, the result is a single-formed factor (unidimensional). Finally, there is only one factor (Infrastructure) that reflects the two variables (public facilities and classrooms). This means that the Infrastructure instrument is proven to have unidimensionality and is empirically valid, so it can be used to develop an evaluation model of character-based Islamic Education learning in Primary Schools.

Construct Validity of Learning Activity

Based on the results of the empirical trial analysis, it is concluded that there is only one factor (Teaching and Learning Activities) reflecting the two variables (Teaching and Learning process and Classroom Management). The instrument of Teaching and Learning Activities is proven to have uni-dimensionality and is empirically valid. Therefore, it can be used to develop an evaluation model of character-based Islamic Education learning in Primary Schools.

Construct Validity of Student Learning Outcomes

Based on the results of the empirical analysis, it is concluded that there is only one factor (Student Learning Outcomes) reflecting the two variables (Knowledge, Attitudes, and Skills). This means the instrument of student learning outcomes is proven to have unidimensionality and is empirically valid. Therefore, it can be used to develop an evaluation model of character-based Islamic Education learning in Elementary Schools.

Discussion

This research was conducted in two stages of testing; the initial trial and the main trial. At the initial trial stage of this study, it was intended to obtain preliminary information about the clarity of the instrument in the developed evaluation model. At this stage, an assessment was carried out by the school manager regarding the evaluation model that has been through previous assessments by experts, namely evaluation experts, and practitioners of Islamic Education in the Delphi technique. Respondents in this test were school principals, Islamic Education teachers, and classroom teachers.

The method employed in this study entailed the use of an evaluation model comprising a meticulously assembled questionnaire and an assessment sheet to determine the readiness of the model to evaluate the quality of a character-based Islamic Education program. The results obtained from this stage are a quantitative score as well as input, suggestions, and criticism from the respondents, which served as material for enhancing the evaluation model. This stage revealed that the EPK-PAI instrument was highly efficacious and feasible to be used to evaluate character-based Islamic Education programs, with an acquisition rate surpassing 81%.

All components and dimensions of learning quality, as determined by practitioners' assessments, quantitatively met the "very good" criterion for evaluating character-based Islamic Education programs. The next stage was the main trial. It was conducted in eight Elementary Schools from May to July 2019 in Kartasura. The elementary schools consisted of SDN of Pucangan 1, SDN Pucangan 5, SDN Kartasura 1, SDN Kartasura 4, SDN Kartasura 6, SDN Ngadirejo 1, SDN Ngadirejo 4, dan SDIT of Al-Kautsar. The participants in the main trial of the study were school organizers and students.

After testing the construct validity in the main trial, the results indicate that there is only one

factor (Teacher Competence) that encompassed the four variables (personality competency, pedagogical competency, professional competency, and social competency). This suggests that the teacher competency instrument has been demonstrated to possess uni-dimensionality and to be empirically valid, thus it can be utilized to formulate an evaluation model for character-based Islamic Education programs in elementary schools.

Similarly, regarding the aspect of teaching and learning activities, the results indicated that there was a single factor (Teaching and Learning Activities) that encompassed two variables (Teaching and Learning Process, and Classroom Management). This implies that the Teaching and Learning Activities instrument has been demonstrated to possess uni-dimensionality and is empirically valid, thus it can be utilized to formulate an evaluation model for character-based Islamic Education learning in primary schools.

The results of this study conclusively prove that the teacher and the teaching and learning process have a crucial role in shaping the character of students through the learning methods employed. A plethora of methods can be applied in schools for character education including habituation method, exemplary method, spontaneous refraction, routine refraction and conditioning, direct and indirect methods, through separate and integrated lessons, activities outside of subjects, advice and attention, as well as reward and punishment. These methods are expected to positively shape the character of students and facilitate the implementation of character education (Marzuki & Hapsari, 2015).

Furthermore, in terms of the construct validity of school facilities and infrastructure, the results indicate that there is only one factor (Infrastructure) reflecting the two variables (Public Infrastructure and Classrooms). This implies that the infrastructure instrument has been demonstrated to possess uni-dimensionality and is empirically valid, thus it can be utilized to formulate an evaluation model for character-based Islamic Education programs in elementary schools.

This is in line with Tusriyanto et al. (2022) who posit that school is a second home for students given that children spend a significant amount of time in the educational environment. Schools are also formal educational institutions that systematically implement guidance, teaching, and training programs to optimize the development of students' potential. Therefore, school facilities and infrastructure play a crucial role in the success of students' character building.

Conclusion

This study has developed an evaluation model of the character-based Islamic Education program (EPK-PAI) in Elementary Schools. The EPK-PAI model was developed through systematic steps adapted from the Borg & Gall model, including information gathering, planning, initial product development, initial field testing, product revision, and field testing. The EPK-PAI model comprises four types of evaluation instruments, which are classified into three dimensions: input, process, and output. Specifically, teacher competence and school infrastructure are used to evaluate the input of the character-based Islamic Education learning program, teaching and learning activities consisting of the learning process and classroom management that can be applied to evaluate the process. Meanwhile, the student learning outcomes in the aspects of knowledge, attitudes, and skills can be applied to evaluate the output of the character-based Islamic Education learning

The EPK-PAI model has construct validity and reliability that can be relied upon and can be utilized to evaluate character-based Islamic education programs in elementary schools. The EPK-PAI model has reliable construct validity and reliability, thereby rendering it suitable for evaluating character-based Islamic education programs in the elementary school context.

Future research endeavors are anticipated to cultivate a new, innovative evaluation model that is adaptable to various educational purposes and levels, and that incorporates a diversity of theoretical frameworks.

Acknowledgment

The researchers would like to acknowledge Sulasmi, S.Pd (Principal of Pucangan Elementary School), and Mr. Nugroho, S.Pd. (Religion Teacher at State Elementary School/SDN 4 of Kartasura. Also, we would like to thank the eights schools as the venue of this's study pilot, i.e. eight Primary Schools in Kartasura, namely State Elementary School (SDN) of Pucangan 1, SDN Pucangan 5, SDN Kartasura 1, SDN Kartasura 4, SDN Kartasura 6, SDN Ngadirejo 1, SDN Ngadirejo 4, dan Integrated Islamic Elementary School (SDIT) of Al-Kautsar.

References

Anidi, A. (2017). An evaluation model of Islamic learning education program in madrasah aliyah. *Jurnal Penelitian dan Evaluasi Pendidikan, 21*(1), 21-33.

Aziz, S., Mahmood, M., & Rehman, Z. (2018). Implementation of CIPP Model for Quality Evaluation at School Level: A Case Study. *Journal of Education and Educational Development, 5*(1), 189-206.

Binks, A. N., Kenway, S. J., Lant, P. A., & Head, B. W. (2016). Understanding Australian household water-related energy use and identifying physical and human characteristics of major end uses. *Journal of Cleaner Production*, *135*, 892-906.

Budimansyah, D. (2011). *Pengantar Pendidikan Kewarganegaraan untuk Membangun Karakter Bangsa*. Kota Samarinda, Indonesia:Widya Aksara Pers.

Chelimsky, E. (1989). *Program Evaluation Pattern and Directions*. Washington DC, USA: Society For Public Administration.

Djaali, P. M. (2000). Ramly. *Pengukuran Dalam Bidang Pendidikan*. Jakarta, Indonesia: PPs UNJ.

Gall, M. D., Borg, W. R., & Gall, J. P. (1996). *Educational research: An introduction*. London, UK: Longman Publishing Group.

Jumadi, J., Sutjipto, S., & Hanafi, I. (2022). The Evaluation of Character Education Program Implementation in Al-Azhar Islamic Elementary School Kelapa Ivory North Jakarta. *International Journal of Environmental, Sustainability, and Social Science, 3*(2), 409-429.

Kirkpatrick, D. L. (1998). *Evaluating training programs: The four levels*. San Francisco, USA: Berrett-Koehler Publishers.

Kementerian Pendidikan Nasional, (2010). *Badan Penelitian dan Pengembangan, Pengembangan Pendidikan Budaya dan Karakter Bangsa*. Jakarta, Indonesia: Kementerian Pendidikan Nasional.

Lickona, T. (2022). *Mendidik untuk membentuk karakter*. Magelang, Indonesia: Bumi Aksara.

Mardapi. D. (2000). *Evaluasi pendidikan. Proceeding of the Konvensi Pendidikan Nasional on 19-23 September 2000*. Jakarta, Indonesia: Universitas Negeri Jakarta.

Majid, A., & Andayani, D. (2004). *Pendidikan agama Islam berbasis kompetensi: konsep dan implementasi kurikulum 2004*. Bandung, Indonesia: Remaja Rosdakarya.

Marzuki, M., & Hapsari, L. (2015). Pembentukan karakter siswa melalui kegiatan kepramukaan di MAN 1 Yogyakarta. *Jurnal Pendidikan Karakter, 6*(2), 142-155.

Meyer, M., Roberts, S. O., Jayaratne, T. E., & Gelman, S. A. (2020). Children's beliefs about causes of human characteristics: Genes, environment, or choice?. *Journal of Experimental Psychology: General*, *149*(10), 1935.

Muhlisin, M., Ekosiswoyo, R., Rahman, M., & Sugiyo, S. (2019). Model of Character Education Based on Ubudiyah Muamalah in Madrasah Aliyah Darul Falah Pati. *The Journal of Educational Development, 7*(2), 78-89.

Musfiroh, T. (2008). *Memilih, Menyusun dan Menyajikan Cerita*. Yogyakarta, Indonesia: Tiara Wacana.

Nurindarwati, R., Mulyoto, M., & Hasyim, D. (2022). Evaluation Model of Character Education Programs in Islamic Religious Education and Moral Education Learning at SMP Negeri I Surakarta. *Journal Research of Social, Science, Economics, and Management, 1*(9), 1453-1472

Plomp, T. (2000). *Development Research in/on Educational Development*. Retrieved from https://hub.hku.hk/bitstream/10722/44041/1/content1.pdf

Ramayulis (2005). *Ilmu Pendidikan Islam*. Jakarta, Indonesia: Kalam Mulia.

Ramli, T. (2003). *Revolusi Akhlak: Pendidikan Karakter*. Yogyakarta, Indonesia: Permata Buku.

Rutman, L. & Mowbray, G. (1983). *Understanding Program Evaluation*. Beverly Hills, American: Sage Publications.

Stufflebeam, D. L., & Shinkfield, A. J. (2012). *Systematic evaluation: A self-instructional guide to theory and practice (Vol. 8)*. Dordrecht, Netherlands: Springer Science & Business Media.

Sudijono, A. (2003). *Pengantar Statistik Pendidikan*. Jakarta, indonesia: Raja Grafindo.

Tusriyanto, T., Lisdiana, A., Purwasih, A., Karsiwan, K., Handayana, S., Baidhowi, M. M., ... & Srinatin, S. (2022). Penerapan Model Pendidikan Ramah Anak Dalam Mendukung Partisipasi Siswa. *Elementary: jurnal ilmiah pendidikan dasar, 8*(1), 51-64.

Widoyoko, E. P., & Mulyatiningsih, M. (2007). Evaluasi Program Pembelajaran. *Yogyakarta: Pustaka Pelajar, 238*.

Wirawan. (2006). *Evaluasi : Teori , Model, Metodologi, Standar, Aplikasi & Profesi*. Jakarta, Indonesia: Rajawali Pers.