

## **Quantitative Evaluation Research and Its Implications for the Development of Islamic Education**

**Fahriah.**

Antasari State Islamic University Banjarmasin, South Kalimantan, Indonesia  
Email: galoehfahri@gmail.com

**Nuril Huda.**

Antasari State Islamic University Banjarmasin, South Kalimantan, Indonesia  
Email: nurilhuda@uin-antasari.ac.id

**Dina Hermina**

Antasari State Islamic University Banjarmasin, South Kalimantan, Indonesia  
Email: dinahermina@uin-antasari.ac.id

**Hamrani**

Hasnur Polytechnic, Banjarmasin, South Kalimantan, Indonesia  
Email: hamrani@polihasnur.ac.id

### **Abstract**

This study aims to describe quantitative evaluation research and its implications for the development of Islamic education in Indonesia. The research was conducted through library research by reviewing literature that discusses educational evaluation models and theories such as CIPP, Goal-Oriented, Goal-Free, and Kirkpatrick, as well as relevant policy documents from educational institutions and the Ministry of Religious Affairs. The analytical approach used is descriptive-analytical, including the identification of key concepts, classification of evaluation approaches, and synthesis of implications for educational practice and policy. The findings indicate that quantitative evaluation research plays a strategic role in systematically and measurably assessing the effectiveness of educational programs and policy interventions. It provides empirical evidence to strengthen institutional accountability, support evidence-based decision-making, and encourage continuous improvement in educational management. In practice, the application of appropriate evaluation models enables Islamic educational institutions to measure the impact of programs such as Madrasah Digital Transformation, Religious Moderation, and Religious Literacy Assessment, thereby strengthening monitoring and follow-up mechanisms.

**Keywords:** Evaluation Research, Quantitative, Islamic Education

### **Abstrak**

Penelitian ini bertujuan untuk mendeskripsikan riset evaluasi kuantitatif dan implikasinya terhadap pengembangan pendidikan Islam di Indonesia. Penelitian dilakukan melalui metode kajian pustaka (library research) dengan penelusuran literatur yang membahas model-model dan teori evaluasi pendidikan seperti CIPP,

Goal-Oriented, Goal-Free, dan Kirkpatrick, serta dokumen kebijakan yang relevan dari institusi pendidikan dan Kementerian Agama. Pendekatan analisis yang dipakai bersifat deskriptif-analitis; langkah analisis meliputi identifikasi konsep kunci, klasifikasi pendekatan evaluatif, dan sintesis implikasi terhadap praktik serta kebijakan pendidikan. Hasil kajian menunjukkan bahwa riset evaluasi kuantitatif memiliki peran strategis dalam menilai efektivitas program pendidikan dan intervensi kebijakan secara sistematis dan terukur. Riset evaluasi menyediakan bukti empiris yang dapat memperkuat akuntabilitas lembaga, mendukung proses pengambilan keputusan berbasis bukti, serta mendorong perbaikan berkelanjutan dalam penyelenggaraan pendidikan. Secara praktis, penerapan model evaluasi yang sesuai memungkinkan lembaga pendidikan Islam untuk mengukur dampak program seperti Transformasi Digital Madrasah, moderasi beragama, serta inisiatif asesmen literasi keagamaan, sehingga memperkuat mekanisme monitoring dan tindak lanjut kebijakan.

**Kata Kunci:** Riset Evaluasi, Kuantitatif, Pendidikan Islam

## A. Introduction

Islamic education in Indonesia continues to undergo transformation in line with the development of science, technology, and the socio-religious dynamics of society. Amidst the acceleration of globalisation and digitalisation, Islamic educational institutions are faced with the challenge of maintaining a balance between Islamic ideals and the demands of 21st-century competencies. The Ministry of Religious Affairs of the Republic of Indonesia emphasises that Islamic education must be the foundation for the emergence of a moderate, tolerant, and integrity-driven generation in facing the challenges of the times.<sup>1</sup> These efforts are implemented through various policies and programmes such as strengthening religious moderation, digital transformation of madrasahs, and developing religious literacy in Islamic educational units.<sup>2</sup>

However, there are still a number of fundamental issues that hinder the effectiveness of these Islamic education policies and programmes. Many of the programmes that have been implemented have not been accompanied by a systematic scientific evaluation mechanism, especially one based on a quantitative approach. Evaluations are often limited to administrative reporting rather than empirical assessments of the results and impact of policies.<sup>3</sup> This condition makes it difficult to measure the extent to which religious programmes actually improve the quality of learning, Islamic character, and the competitiveness of students. For example, the implementation of the National Assessment of Basic Religious Literacy (ANLDB) held in 2025 has the potential to be an important source of data for evaluation research to assess the effectiveness of the implementation of religious literacy policies in schools and madrasahs. This programme still needs to be strengthened in terms of quantitative data analysis so that the mapping results can be

---

<sup>1</sup> Kementerian Agama Republik Indonesia, *Moderasi Beragama: Kunci Indonesia Emas 2045 – Penjelasan Direktur Diktis*, diakses 10 Oktober 2025, <https://pendis.kemenag.go.id/direktorat-perguruan-tinggi-keagamaan-islam/moderasi-beragama-kunci-indonesia-emas-2045-ini-penjelasan-direktur-diktis>.

<sup>2</sup> Direktorat Jenderal Pendidikan Islam, *Transformasi Digital Madrasah 2024: Arah Kebijakan dan Implementasi*, Kementerian Agama RI, diakses 11 Oktober 2025, <https://madrasah.kemenag.go.id/berita/transformasi-digital-madrasah>

<sup>3</sup> Muhammad Zuhdi, "Evaluasi Kebijakan Pendidikan Islam: Tantangan dan Prospek," *Jurnal Pendidikan Islam* 13, no. 2 (2023), h. 221–240.

linked to indicators of religious moderation practices and improvements in learning quality.<sup>4</sup>

In this context, quantitative evaluation research has strategic urgency. Evaluation research plays a role in scientifically assessing the value, benefits, effectiveness, and sustainability of the implementation of Islamic education policies. The quantitative approach allows for the systematic collection and analysis of numerical data, thus providing reliable empirical evidence for decision making. The results of this evaluative research form the basis for data-driven recommendations for curriculum development, teacher capacity building, and monitoring the implementation of Islamic education programmes.

Thus, the development of quantitative evaluation research in Islamic education is an important instrument to ensure that Islamic values are integrated effectively, measurably, and in line with the demands of educational modernisation.

## B. Literature Review

Evaluation research is an important discipline in educational research that aims to systematically assess the effectiveness, efficiency, relevance, and sustainability of a programme. According to Daniel L. Stufflebeam, evaluation is the process of providing useful information for decision-makers to determine the value and quality of an educational programme.<sup>5</sup> This approach is very important in the context of modern education because it provides an empirical basis for objectively assessing the success of policies and programmes. In quantitative evaluation research, data is collected through standardised instruments so that analysis can be carried out statistically, making the research results replicable, valid, and reliable.<sup>6</sup>

Conceptually, quantitative evaluation research is rooted in the positivist paradigm that emphasises objectivity, measurement, and analysis of relationships between variables.<sup>7</sup> Creswell explains that the quantitative approach is used to test theories by examining the relationships between variables through systematic measurement using instruments that produce numerical data.<sup>8</sup> This allows researchers to assess the extent to which educational programme objectives have been achieved and how certain factors influence their effectiveness. In the context of Islamic education, the quantitative approach helps educational institutions assess the effectiveness of policies such as religious moderation, digital transformation of madrasas, and Islamic Religious Education (PAI) teacher training, through measurable indicators and instruments.<sup>9</sup>

Meanwhile, validity and reliability are fundamental components in quantitative evaluation research. Internal validity relates to the extent to which observed changes in outcomes can be trusted as a direct result of the intervention being evaluated,

---

<sup>4</sup> Detik News, “Kemenag Gelar Asesmen Nasional Literasi Keagamaan 2025 bagi Guru dan Siswa SD”, diakses 16 Oktober 2025, <https://www.detik.com/hikmah/khazanah/d-8191892/kemenag-gelar-asesmen-nasional-literasi-keagamaan-2025-bagi-guru-dan-siswa-sd>.

<sup>5</sup> Daniel L. Stufflebeam dan Anthony J. Coryn, *Evaluation Theory, Models, and Applications*, 2nd ed. (San Francisco: Jossey-Bass, 2014), h. 12-14.

<sup>6</sup> Daniel L. Stufflebeam dan Anthony J. Coryn, *Evaluation Theory ...*, h.15.

<sup>7</sup> John W. Creswell, *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*, 5th ed. (Thousand Oaks: SAGE Publications, 2018), h. 20–21.

<sup>8</sup> John W. Creswell, *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research*, 6th ed. (Boston: Pearson, 2023), h. 32–34.

<sup>9</sup> M. Alamsyah, “Analisis Efektivitas Asesmen Kuantitatif dalam Meningkatkan Pendidikan Agama Islam,” *Journal of Education Research and Policy* 4, no. 2 (2023): h. 45–46.

rather than other factors.<sup>10</sup> Gary W. Moore emphasises that the quality of evaluative research is largely determined by the accuracy of the measuring instruments and control over external variables that may influence the data.<sup>11</sup> Thus, quantitative evaluation research must be carefully designed to ensure that the conclusions drawn truly reflect the empirical conditions in the field.

Evaluation models are an important framework for designing and interpreting quantitative evaluation research. One of the most influential models is the CIPP (Context, Input, Process, Product) model developed by Stufflebeam.<sup>12</sup> This model is widely used in the evaluation of Islamic education because it provides a comprehensive overview of programme needs, implementation, and results achieved. Recent research by Alanshori et al. (2024) confirms that the CIPP model is effective for assessing various education programmes in Indonesia, including religious-based programmes.<sup>13</sup> Apart from CIPP, Kirkpatrick's four-level model (reaction, learning, behaviour, results) is also widely used in evaluating PAI teacher training or programmes to improve the competence of educators in madrasahs.<sup>14</sup> Both models are relevant not only in the context of general education but also in the development of Islamic education policies and programmes.

In addition, other evaluation approaches such as Goal-Oriented Evaluation (Tyler) and Goal-Free Evaluation (Scriven) also provide important perspectives in evaluative research. Tyler emphasises the importance of assessing programme goal achievement based on measurable indicators, while Scriven offers a more open alternative by assessing actual impact without being bound by formal goals.<sup>15</sup> The Goal-Free approach has proven relevant in the context of Islamic education, especially when researchers want to uncover the unexpected impacts of programmes such as religious moderation or religious extracurricular activities.<sup>16</sup>

The evaluation framework in Islamic education is also reinforced by classical and contemporary Islamic literature. Bahiyah et al. (2024) explain that evaluation from the perspective of the Qur'an and hadith emphasises the importance of fair, objective, and comprehensive assessment of students' behaviour and learning outcomes.<sup>17</sup> Principles such as muraqabah (supervision), hisbah (moral assessment), and amanah (accountability) are fundamental values that are relevant in formulating a modern Islamic education evaluation model. In another study, Tsani (2023) developed an affective PAI evaluation model that emphasises the integration of spiritual values and behaviour through a quantitative approach based on moral

---

<sup>10</sup> Creswell, *Educational Research*,...h. 102–103.

<sup>11</sup> Gary W. Moore, *Developing and Evaluating Educational Research* (Boston: Little, Brown and Company, 1983), h. 50–52.

<sup>12</sup> Stufflebeam and Coryn, *Evaluation Theory*, ... h. 35–36.

<sup>13</sup> M. Zainuddin Alanshori et al., "The Application of the CIPP Evaluation Model in Educational Programs in Indonesia," *Journal of Educational Evaluation* (2024): h. 6–7.

<sup>14</sup> Donald L. Kirkpatrick dan James D. Kirkpatrick, *Evaluating Training Programs: The Four Levels*, 3rd ed. (San Francisco: Berrett-Koehler, 2006), h. 21–25.

<sup>15</sup> Ralph W. Tyler, *Basic Principles of Curriculum and Instruction* (Chicago: University of Chicago Press, 1949), h. 30–32 ;

Michael Scriven, *Evaluation Thesaurus*, 4th ed. (Newbury Park: SAGE Publications, 1991), h. 10–11.

<sup>16</sup> Scriven, *Evaluation Thesaurus*,... h. 12.

<sup>17</sup> Khoridatul Bahiyah et al., "Learning Evaluation in Islam: Review of the Concept of the Qur'an and Hadith," *Modern Islamic Studies & Civilization* 5, no. 1 (2024): h. 88–89.

indicators.<sup>18</sup> This shows the development of a more contextual evaluation model that meets the needs of Islamic educational institutions today.

Furthermore, recent research shows that quantitative evaluation research has been significantly applied in Islamic educational institutions in Indonesia. Alamsyah's (2023) study, for example, shows that quantitative assessment is effective in improving the quality of PAI learning through the use of standardised instruments and statistical analysis.<sup>19</sup> Meanwhile, Aprianti's (2024) research, which uses the CIPP model in PAI courses at universities, shows that a comprehensive evaluative approach helps improve the quality of learning and institutional accountability.<sup>20</sup> These studies support the idea that quantitative evaluation research not only provides an overview of programme objectives achievement but also serves as an important basis for evidence-based policy development in Islamic education.

In addition to understanding research models and findings, it is also important to distinguish between quantitative and qualitative approaches in evaluation research. Quantitative approaches provide objective measurements of programme effectiveness, while qualitative approaches focus on meaning, experience, and socioreligious contexts that cannot be reduced to numbers.<sup>21</sup> The qualitative approach is important in assessing the process of internalising the values of religious moderation, religious attitudes, or teachers' perceptions of certain policies.<sup>22</sup> In many Islamic education studies, the integration of both approaches through mixed methods has proven to be more comprehensive. Tashakkori and Teddlie emphasise that the mixed approach produces a more complete picture because it combines the strengths of numerical data and the depth of narrative data.<sup>23</sup> In evaluating the Moderate Religion programme, for example, quantitative surveys can show an increase in tolerance, while qualitative interviews can explain the factors that influence this change.<sup>24</sup>

In the context of Islamic education evaluation research, quantitative and qualitative approaches should not be positioned as two opposing methods, but as complementary approaches. The quantitative approach provides an objective picture through standardised measurements, while the qualitative approach provides depth of meaning, context, and an understanding of processes that cannot be reduced to numbers.<sup>25</sup> When combined, the two approaches can produce a more comprehensive evaluation, especially in assessing the cognitive, affective, psychomotor, and spiritual dimensions of Islamic education.<sup>26</sup> Through statistical analysis, researchers can identify the level of achievement of religious learning objectives; while through interviews or observations, researchers can understand the

---

<sup>18</sup> I. Tsani, "Affective Evaluation Model for Islamic Religious Education," *Didaktika* 12, no. 2 (2023): h. 112–114.

<sup>19</sup> Alamsyah, "Analisis Efektivitas Asesmen Kuantitatif..." h. 47.

<sup>20</sup> N. Aprianti, "Evaluation Using the CIPP Model in Islamic Religious Education," *Conciencia* 13, no. 1 (2024): h. 55–57.

<sup>21</sup> Creswell, *Educational Research*, h. 40–42.

<sup>22</sup> Matthew B. Miles, A. Michael Huberman, and Johnny Saldaña, *Qualitative Data Analysis*, 4th ed. (Thousand Oaks: SAGE Publications, 2020), h. 12–15.

<sup>23</sup> Abbas Tashakkori and Charles Teddlie, *Mixed Methodology: Combining Qualitative and Quantitative Approaches* (Thousand Oaks: SAGE Publications, 1998), h. 9–12.

<sup>24</sup> Tsani, "Affective Evaluation Model..." h. 113–114.

<sup>25</sup> Creswell, *Educational Research...* h. 40–42.

<sup>26</sup> Matthew B. Miles, A. Michael Huberman, and Johnny Saldaña, *Qualitative Data Analysis...* h. 12–15.

dynamics of internalisation of values, religious motivation, and student responses to religious material.<sup>27</sup> Therefore, the integration of quantitative and qualitative methods (mixed methods) is a highly strategic approach for Islamic educational institutions to obtain a more comprehensive, accurate, and relevant evaluation of the reality in the field.<sup>28</sup>

### C. Method

This article was written using a qualitative approach with a library research method. Library research was chosen because the purpose of this study was not to collect new empirical data, but rather to analyse and synthesise existing theories and scientific findings on evaluation research in the context of quantitative educational research and its relevance to the development of Islamic education.<sup>29</sup>

Data was collected through literature searches sourced from research methodology textbooks, scientific journal articles, and relevant Ministry of Religious Affairs policy documents. The main sources used as references included Gary W. Moore's work on *Developing and Evaluating Educational Research*, which explains the basic principles of evaluation in educational research,<sup>30</sup> and Muhammad Ali's book entitled *Understanding Behavioural and Social Research*, which provides a theoretical foundation for understanding the characteristics of quantitative and evaluative research in the field of education and social behaviour.<sup>31</sup> In addition, references from Stufflebeam and Coryn (2014), which describe educational evaluation models such as Context–Input– Process–Product (CIPP), Donald L. Kirkpatrick (2006), which explains the four levels of training evaluation (reaction, learning, behaviour, results), and others, were also used.<sup>32</sup>

The analysis procedure was carried out using a descriptive-analytical approach, namely by describing the main theories of evaluation research and analysing their relevance to the development of Islamic education. The analysis was carried out in three stages:

1. Identification of key concepts related to educational evaluation research from the main literature.
2. Classification of approaches and models that can be applied in the context of Islamic education.

Synthesis analysis to draw implications from evaluation research for the development of Islamic education programmes and policies in Indonesia.

### D. Results and Discussion

#### I. Basic Concepts of Evaluation Research in Education

---

<sup>27</sup> I. Tsani, "Affective Evaluation Model...", h. 113–114.

<sup>28</sup> Abbas Tashakkori and Charles Teddlie, *Mixed Methodology: Combining Qualitative and Quantitative Approaches* (Thousand Oaks: SAGE Publications, 1998), h. 9–12

<sup>29</sup> Creswell, *Research Design*, h. 30–32.

<sup>30</sup> Gary W. Moore, *Developing and Evaluating Educational Research* (Boston: Little, Brown and Company, 1983), h. 10–12.

<sup>31</sup> Muhammad Ali, *Memahami Riset Perilaku dan Sosial* (Bandung: Pustaka Cendekia Utama, 2011), h. 45–46

<sup>32</sup> Daniel L. Stufflebeam dan Anthony J. Coryn, *Evaluation Theory, Models, and Applications*, 2nd ed....; Donald L. Kirkpatrick dan James D. Kirkpatrick, *Evaluating Training Programs: The Four Levels*, 3rd ed....

Evaluation research is a branch of educational research that focuses on assessing the value, utility, and effectiveness of a programme or policy.<sup>33</sup> Unlike descriptive or experimental research, which focuses on discovering new facts, evaluation research aims to determine the extent to which a programme has achieved its objectives and to provide a scientific basis for decision-making.<sup>34</sup> Thus, evaluation research is not only academic in nature, but also applicable because its results are used to improve, develop, or discontinue an ongoing education policy.

In the context of education, evaluation research has a strategic function in ensuring quality assurance, accountability, and effectiveness in programme implementation. Gary W. Moore (1983) emphasises that educational evaluation must be carried out systematically through a process of collecting and analysing valid and reliable data to assess the quality of programmes or learning activities<sup>35</sup> Quantitative data-based evaluation enables educational institutions to objectively assess learning outcomes, teacher performance, and policy achievements. This is in line with Daniel L. Stufflebeam's (2014) view, which emphasises that evaluation research is a scientific process that not only assesses results (products) but also the context, inputs, and implementation process of programmes.<sup>36</sup>

Furthermore, Muhammad Ali (2011) explains that evaluation research occupies an important position in behavioural and social research because it serves as a tool for assessing the effectiveness of educational interventions, including in the religious field.<sup>37</sup> The quantitative approach in evaluation research allows researchers to measure the relationship between variables, for example, between the level of teacher participation in training and the improvement of pedagogical competence, or between the implementation of religious moderation policies and the tolerance attitudes of students. Thus, quantitative evaluation research plays a role in linking policies with measurable real-world outcomes.

In the context of Islamic education, evaluation research is an important instrument to ensure that the programmes and policies implemented truly reflect substantive Islamic values, such as justice, responsibility, and benefit. Scientific evaluation can help Islamic educational institutions understand the effectiveness of policies, such as programmes to strengthen religious moderation, the digitisation of madrasas, or the improvement of Islamic Religious Education (PAI) teacher competencies. Thus, evaluation research serves not only as an administrative measurement tool, but also as an epistemological and ethical instrument in maintaining the relevance of Islamic education in the modern era.

Furthermore, evaluation research encourages a paradigm shift from procedure-based education to evidence-based education<sup>38</sup> This approach requires Islamic educational institutions to make decisions based on scientific evaluation results, not merely on assumptions or administrative policies. By integrating evaluation research into the Islamic education system, it is hoped that the policies

---

<sup>33</sup> Peter H. Rossi, Howard E. Freeman, and Mark W. Lipsey, *Evaluation: A Systematic Approach*, 7th ed. (Thousand Oaks: SAGE Publications, 2004), h. 16.

<sup>34</sup> John W. Creswell, *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research*, 6th ed. (Boston: Pearson, 2023), h. 45–46.

<sup>35</sup> Gary W. Moore, *Developing and Evaluating Educational Research*..., h. 55–57.

<sup>36</sup> Daniel L. Stufflebeam and Anthony J. Coryn, *Evaluation Theory, Models, and Applications*..., h. 85–86.

<sup>37</sup> Muhammad Ali, *Memahami Riset Perilaku dan Sosial*..., h. 68–69.

<sup>38</sup> Thomas Guskey, "Data-Driven Decision Making and the Future of Educational Evaluation," *Educational Measurement: Issues and Practice* 42, no. 1 (2023), h. 12–19.

and programmes developed will be more effective, transparent, and oriented towards comprehensively improving the quality of students — both intellectually and spiritually.

## **2. Approaches and Models of Evaluation Research and Their Application in Islamic Education**

### **a. Evaluation Research Approach**

Evaluation research in the context of education has various approaches and models used to assess the effectiveness of programmes, policies, and learning activities. In general, there are three main approaches that can be used, namely qualitative, quantitative, and mixed methods. The quantitative approach is a method that is often used because it provides objective results through numerical data and statistical analysis.<sup>39</sup> According to John W. Creswell, the quantitative approach is used to test theories by examining the relationship between variables using instruments that produce numerical data that can be analysed statistically.<sup>40</sup> Thus, quantitative evaluation research allows researchers to assess the extent to which a programme has achieved its objectives based on a comparison between actual results and predetermined criteria.

In practice, quantitative evaluation research can use various designs, such as pure experiments, quasi-experiments, pretest–posttest, or evaluative surveys.<sup>41</sup> For example, to assess the effectiveness of a PAI learning model based on religious moderation values, researchers can use a pretest–posttest control group design to measure changes in tolerance attitudes before and after the programme intervention. This approach emphasises the principles of objectivity, validity, and replication — where research results can be retested to ensure their reliability. Gary W. Moore emphasises that reliability and validity are two key components in evaluation research, because conclusions can only be trusted if the measuring instruments truly reflect the objectives to be evaluated.<sup>42</sup>

In the context of Islamic education, a mixed methods approach is sometimes necessary. This approach is used when dimensions of values, beliefs, or religious behaviour are difficult to measure with numbers alone. Tashakkori and Teddlie explain that the integration of quantitative and qualitative data provides a more comprehensive understanding of the social and religious context of education.<sup>43</sup> For example, research on the effectiveness of religious moderation policies does not only assess the increase in tolerance scores numerically, but also analyses how these values are internalised in learning practices. Thus, a mixed approach helps researchers see the integration between empirical data and normative meanings in Islamic education.

### **b. Evaluation Research Models**

Various evaluation research models have been developed to provide a framework for assessing the success of educational programmes. Some of the most

---

<sup>39</sup> Creswell, *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches...*, h. 13

<sup>40</sup> Creswell, *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches...*, h. 20–21

<sup>41</sup> Donald Ary et al., *Introduction to Research in Education*, 8th ed. (Belmont: Wadsworth, 2010), h. 302.

<sup>42</sup> Gary W. Moore, *Developing and Evaluating Educational Research...*, h. 95.

<sup>43</sup> Abbas Tashakkori and Charles Teddlie, *Mixed Methodology: Combining Qualitative and Quantitative Approaches* (Thousand Oaks: SAGE Publications, 1998), h. 12–14.

relevant models in Islamic education research include the CIPP Model, Kirkpatrick Model, Goal-Oriented Evaluation, and Goal-Free Evaluation.

### 1. CIPP Model (Context–Input–Process–Product)

The CIPP model developed by Daniel L. Stufflebeam is one of the most comprehensive models in evaluation research.<sup>44</sup> This model views evaluation as a systematic process to assist decision-makers by assessing four main components:

1. Context Evaluation, to identify the needs, problems, and background of the programme.
2. Input Evaluation, to assess strategies, resources, and implementation plans.
3. Process Evaluation, to monitor the implementation of the programme.
4. Product Evaluation, to assess the results, impact, and sustainability of the programme.

In the context of Islamic education, the CIPP model is highly relevant for evaluating national-scale policies such as Madrasah Digital Transformation. Context evaluation can be conducted by assessing the needs and readiness of madrasahs for ICT infrastructure through a digital readiness survey. Furthermore, input evaluation involves assessing resource support, teacher training, and the availability of digital learning devices. Process evaluation assesses the implementation of online learning activities through structured observation, while product evaluation analyses improvements in learning outcomes and student satisfaction using pretest–posttest tests. The quantitative data obtained from these four stages is then statistically analysed to determine the program's success rate based on predefined effectiveness indicators.

### 2. Kirkpatrick Model

The four-level model developed by Donald L. Kirkpatrick is widely used in training and human resource development evaluations.<sup>45</sup> This model assesses programme effectiveness through four levels: reaction, learning, behaviour, and results. In Islamic education, this model is very useful for evaluating PAI teacher training programmes or capacity building for madrasah educators.

For example, teacher training evaluation can be conducted by:

- Reaction – measuring participant satisfaction with the content and facilitators of the training through a Likert scale questionnaire.
- Learning – assessing improvements in Islamic pedagogical knowledge using objective tests (pre-test–post-test).
- Behaviour – observing changes in teaching behaviour after training using an observation rubric.
- Results – assessing the long-term impact on learning quality and student learning outcomes through average score analysis.

This quantitative approach produces data that can be tested using ANOVA or paired-sample t-test analysis to determine the effectiveness of the programme.

### 3. Goal-Oriented Evaluation Model

This model was developed by Ralph W. Tyler, who emphasised that evaluation should focus on the achievement of programme objectives.<sup>46</sup> In this

---

<sup>44</sup> Daniel L. Stufflebeam and Anthony J. Coryn, *Evaluation Theory, Models, and Applications...*, h. 23–25.

<sup>45</sup> Donald L. Kirkpatrick and James D. Kirkpatrick, *Evaluating Training Programs: The Four Levels...*, h.21–25.

<sup>46</sup> Ralph W. Tyler, *Basic Principles of Curriculum and Instruction* (Chicago: University of Chicago Press, 1949), h. 104.

approach, the success of a programme is measured by comparing the objectives that have been formulated with the actual results obtained. For example, the evaluation of the Islamic Religious Education (PAI) curriculum policy in madrasahs.

Researchers can set goals—such as improving students' religious literacy and moderate attitudes—and then develop measurable indicators such as ANLDB scores and tolerance attitude scales. The evaluation is carried out by comparing pretest and posttest scores after the implementation of a new policy or curriculum. The results can show the effectiveness of policy implementation in achieving the expected Islamic education goals.

#### 4. Goal-Free Evaluation Model

Unlike Tyler, Michael Scriven introduced the Goal-Free Evaluation model, which assesses actual results without referring to the formal objectives of the programme.<sup>47</sup> This approach seeks to avoid evaluation bias that arises from an excessive focus on administrative targets. In the context of Islamic education, this model can reveal unexpected effects of religious programmes. For example, in the Religious Moderation programme in madrasahs, research can identify changes in students' social behaviour—such as increased intergroup empathy or participation in interfaith activities—even though these indicators are not written in the programme objectives. Data is collected through behavioural surveys using a Likert scale, then analysed using descriptive and correlational statistics to identify trends in attitude change. This approach helps ensure that evaluation is not limited to administrative objectives, but also reveals broader socio-cultural impacts.

#### c. Comparison of Evaluation Models

Each evaluation research model has advantages and limitations depending on the context of its use. The CIPP model excels in the completeness of its analysis because it covers all stages of the programme (from planning to results), but its implementation requires considerable resources and time. The Kirkpatrick model is simpler and more effective for evaluating short-term training, but is less adequate for macro policies. The Goal-Oriented model is effective for assessing the achievement of specific objectives, but tends to ignore unexpected results. Meanwhile, the GoalFree model is more flexible and open to new findings, but can lose direction if not accompanied by clear assessment indicators.<sup>48</sup>

In practice, Islamic education researchers often combine two or more models to make the evaluation results more comprehensive. For example, the evaluation of Religious Moderation can combine the CIPP Model to assess context and process with the Goal-Free Model to assess unplanned social impacts.

This kind of integrative approach allows researchers to assess the effectiveness of programmes not only in terms of administrative success, but also in terms of spiritual, social, and religious values. Comprehensive evaluation research can strengthen evidence-based education culture, improve the accountability of Islamic educational institutions, and ensure that every policy is truly oriented towards improving the quality and Islamic character of students.

The application of quantitative evaluation research in Islamic education can be seen from various empirical studies in Indonesia that demonstrate the effectiveness of this method in assessing and improving the quality of learning. One important study was conducted by Alamsyah (2023), who evaluated the effectiveness of

<sup>47</sup> Michael Scriven, *Evaluation Thesaurus...*, h. 57.

<sup>48</sup> Daniel L. Stufflebeam and Anthony J. Coryn, *Evaluation Theory, Models, and Applications...*, h. 44–45.

quantitative assessment in Islamic Religious Education (PAI) learning in primary schools. The results of his research show that the use of standardised instruments based on the Likert scale increases the accuracy of teachers' assessments in measuring students' religious competence.<sup>49</sup> These standardised instruments also help schools map learning outcomes more systematically, so that the learning improvement process can be carried out based on valid data.

Another study by Alanshori et al. (2024) applied the CIPP evaluation model to assess the effectiveness of the PAI teacher quality improvement programme. The results showed that the process and product components contributed most significantly to improving teachers' pedagogical competence.<sup>50</sup> CIPP-based quantitative evaluation helps educational institutions identify aspects that need improvement, such as the appropriateness of teaching methods, the effectiveness of training, and student learning outcomes. These findings reinforce the relevance of number-based evaluation models in improving the professional quality of PAI teachers.

Another example can be seen in Aprianti's (2024) research, which evaluated the effectiveness of PAI courses in higher education using the CIPP model. The results showed that the context and input component scores were in the good category, but the process component required intervention due to a lack of learning innovation and digital technology utilisation.<sup>51</sup> Through quantitative evaluation, study programmes can make evidence-based decisions to improve curriculum design, learning facilities, and lecturer quality.

In addition, Tsani's (2023) research on affective evaluation models in Islamic Religious Education shows how quantitative approaches can be used to measure affective aspects such as religious attitudes, tolerance, and student morals through tested psychometric scales.<sup>52</sup> These findings indicate that religious values—which have long been considered difficult to measure—can actually be evaluated objectively through valid and reliable quantitative measurement tools.

From these various empirical examples, it is evident that quantitative evaluation research is not only theoretically relevant but also has a significant practical impact on improving the quality of Islamic education. The use of standardised instruments, comprehensive evaluation models such as CIPP, and statistical analysis techniques can provide a clear picture of programme effectiveness, learning quality, and educator competence. These empirical findings reinforce the argument that quantitative evaluation is a strategic approach to supporting evidence-based decision-making in Islamic education.

### **3. Objectives and Directions of Evaluation Research in Islamic Education**

Evaluation research has a strategic direction in ensuring the quality, effectiveness, and relevance of Islamic education programmes. In the context of education policy in Indonesia, evaluation research is necessary to ensure that every programme and policy truly achieves its objectives, has a measurable impact, and

---

<sup>49</sup> M. Alamsyah, "Analisis Efektivitas Asesmen Kuantitatif dalam Meningkatkan Pendidikan Agama Islam," ... h. 45–47

<sup>50</sup> M. Zainuddin Alanshori et al., "The Application of the CIPP Evaluation Model in Educational Programs in Indonesia," ... h. 6–7.

<sup>51</sup> N. Aprianti, "Evaluation Using the CIPP Model in Islamic Religious Education," ... h. 55–57.

<sup>52</sup> I. Tsani, "Affective Evaluation Model," ... h. 112–114.

provides tangible benefits for students and the community.<sup>53</sup> Research-based evaluation helps minimise speculative decisions, as it is based on empirical data that shows the extent to which Islamic education policies are effective.

From an Islamic education perspective, the direction of evaluation research is not solely oriented towards academic achievement, but also towards the moral and spiritual dimensions of students. This is in line with the principles of ta'dib and tarbiyah, which place education as a process of shaping the whole personality.<sup>54</sup> Thus, evaluation research in Islamic education aims not only to measure the success of programmes, but also to ensure that the educational process truly produces individuals who are knowledgeable, have good character, and contribute to society.

In the context of the Ministry of Religious Affairs' policies, there are programmes that are relevant for assessment through evaluation research, for example: (1) the Religious Moderation Programme, (2) Madrasah Digital Transformation, and (3) the National Assessment of Basic Religious Literacy (ANLDB). All three have different orientations, but they all require measurable evaluation to determine their effectiveness and direction for improvement.

#### **a. Evaluation of the Religious Moderation Programme**

The Religious Moderation Programme is one of the Ministry of Religious Affairs' strategic policies aimed at instilling values of tolerance, balance, and national commitment in educational institutions.<sup>55</sup> The purpose of the evaluation research on this programme is to assess its effectiveness in shaping the religious behaviour of students and teachers in line with the principles of Islam wasathiyah.

The evaluation research can be directed to measure:

1. The effectiveness of the moderation learning programme, through analysis of students' knowledge and understanding of the concept of religious moderation;
2. Changes in social attitudes and behaviour, by assessing the extent to which students demonstrate tolerance and interfaith empathy after participating in the programme;
3. Institutional readiness and support, such as the involvement of madrasah principals, Islamic education teachers, and the role of the community in strengthening the value of moderation.

Such evaluative findings are important because they can show whether religious moderation policies have had an impact on the transformation of values and behaviour, or whether they have remained at the level of socialisation and ceremony.

#### **b. Evaluation of the Madrasah Digital Transformation Programme**

The main objective of the evaluation research on the Madrasah Digital Transformation programme is to ensure that the digitisation of learning not only improves technical aspects but also strengthens the quality of the learning process and character building of students.<sup>56</sup> Evaluation is needed to assess the effectiveness of teacher training, infrastructure readiness, and the impact of technology use on student learning outcomes and participation.

---

<sup>53</sup> Egon G. Guba dan Daniel L. Stufflebeam, *Evaluation in Education: The Need and the Response* (Bloomington: Phi Delta Kappa, 1970), h. 14.

<sup>54</sup> Syed Muhammad Naquib al-Attas, *The Concept of Education in Islam* (Kuala Lumpur: ISTAC, 1991), h. 45–46

<sup>55</sup> Kementerian Agama RI, *Moderasi Beragama: Buku Saku Penguatan Moderasi Beragama* (Jakarta: Kemenag RI, 2022), h. 5–7.

<sup>56</sup> Kementerian Agama RI, *Roadmap Transformasi Digital Madrasah 2023–2029* (Jakarta: Direktorat KSKK Madrasah, 2023), h. 12–13.

A quantitative approach in evaluation research can help measure the achievement of the following indicators:

1. Madrasah readiness for technology, such as internet access, digital learning devices, and technical support;
2. Teachers' competence in the use of learning technology, which can be measured through pre- and post-training assessments (pretest–posttest);
3. The impact on student learning outcomes and motivation, through perception surveys and assessment result analysis.

The results of the evaluation research will show the extent to which digital transformation actually contributes to the quality of learning. If disparities are found between madrasahs—for example, due to limited facilities or teacher competence—the evaluation data can be used as a basis for more targeted follow-up policy planning.

#### **c. Evaluation of the National Assessment of Basic Religious Literacy (ANLDB)**

The National Assessment of Basic Religious Literacy (ANLDB) programme developed by the Ministry of Religious Affairs' Research, Development and Training Agency aims to map the religious literacy abilities of madrasah students in Indonesia.<sup>57</sup> The purpose of the evaluation research on ANLDB is to assess the extent to which the assessment instruments are able to describe students' religious literacy competencies, as well as how the assessment results are used to improve education policy.

The evaluation research can be directed towards the following objectives:

1. Assessing the reliability and validity of assessment instruments, so that ANLDB data truly describes students' religious literacy abilities objectively;
2. Analysing the relationship between religious literacy and moderation, through a correlational approach to test the influence of literacy levels on socioreligious behaviour;
3. Assessing the effectiveness of policy follow-up, namely the extent to which assessment results are used to improve the curriculum, teacher training, and religious learning policies.

With systematic evaluation research, ANLDB not only functions as a measurement tool, but also as a basis for decision-making that can strengthen the quality and direction of national Islamic education.

#### **4. Implications of Evaluation Research for Islamic Education**

Evaluation research in Islamic education has broad implications for policy development, improving the quality of education, and shaping the character of the nation based on Islamic values. Evaluations conducted systematically and based on data not only serve to measure the success of programmes but also form the foundation for evidence-based decision making in Islamic educational institutions.<sup>58</sup> Thus, the results of evaluation research can strengthen the position of Islamic education as a system that is adaptive to changing times without losing its spiritual and moral orientation.

##### **a. Implications for Improving Quality and Accountability in Education**

---

<sup>57</sup> Badan Litbang dan Diklat Kemenag RI, *Pedoman Asesmen Nasional Literasi Dasar Keagamaan (ANLDB)* (Jakarta: Puslitbang Pendidikan Agama dan Keagamaan, 2024), h. 3–5

<sup>58</sup> Egon G. Guba dan Daniel L. Stufflebeam, *Evaluation in Education: The Need and the Response...*, h. 14.

The quantitative evaluation research approach has important implications for improving the quality and accountability of education, especially in the context of Islamic education. Through the use of standardised instruments, educational institutions can obtain more accurate data on student learning outcomes, the effectiveness of teaching methods, and the success of educational programmes. Quantitative data enables institutions to identify specific aspects of learning that need improvement, not based on subjective assumptions, but through empirically verifiable findings.<sup>59</sup> Alamsyah's (2023) findings, for example, show that the application of quantitative assessment in PAI can improve teachers' ability to conduct more objective assessments focused on measurable competency indicators.<sup>60</sup>

In terms of accountability, evaluation models such as CIPP provide a comprehensive framework for assessing educational programmes as a whole, enabling educational institutions to demonstrate transparency in their performance to stakeholders such as the government, parents, and the community.<sup>61</sup> The results of research by Alanshori et al. (2024) confirm that the use of the CIPP evaluation model encourages madrasas and schools to report their performance more systematically, especially in terms of the learning process and outcomes.<sup>62</sup> Thus, quantitative evaluation research not only helps improve quality but also strengthens a data-based culture of accountability in Islamic educational institutions.

In addition, quantitative evaluation research plays a strategic role in supporting effective decision-making. Valid and reliable numerical data makes it easier for leaders of educational institutions to conduct benchmarking, predict the need for teacher competency improvement, and plan targeted interventions.<sup>63</sup> This is important considering that Islamic educational institutions, including madrasahs and Islamic boarding schools, are facing demands for modernisation, digitalisation of learning, and improvement of national quality standards. Quantitative evaluation provides a strong foundation for assessing whether these innovation programmes are effective and have a significant impact on improving the quality of education.

Thus, the main implications of quantitative evaluation research on improving the quality and accountability of Islamic education are the availability of strong empirical evidence for continuous improvement, increased teacher professionalism, and the strengthening of a more transparent and measurable performance reporting system. Quantitative evaluation helps educational institutions move from an administrative culture to a data-driven quality culture, which is an important foundation for strengthening the quality of Islamic education in the modern era.

#### **b. Implications for Strengthening Islamic Education Policy**

Quantitative evaluation research contributes significantly to strengthening Islamic education policy, particularly by providing objective data that forms the basis for policy formulation and revision. The quantitative approach enables policymakers to map the effectiveness of programmes more accurately, for example, in policies to improve madrasah quality, develop Islamic education teachers' professionalism, or strengthen religious moderation programmes in the educational environment. With

---

<sup>59</sup> Creswell, *Educational Research*,... h. 32–34.

<sup>60</sup> M. Alamsyah, "Analisis Efektivitas Asesmen Kuantitatif dalam Meningkatkan Pendidikan Agama Islam,"... h. 45–47.

<sup>61</sup> Daniel L. Stufflebeam and Anthony J. Coryn, *Evaluation Theory, Models, and Applications*,... h. 35–36.

<sup>62</sup> M. Zainuddin Alanshori et al., "The Application of the CIPP Evaluation Model in Educational Programs in Indonesia," ... h. 6–7.

<sup>63</sup> Gary W. Moore, *Developing and Evaluating Educational Research*,... h. 50–52

statistical data on learning outcomes, teacher performance, and training effectiveness, the government can identify which parts of the policy are working optimally, which require intervention, and which should be removed because they do not have a significant impact.

In addition, quantitative evaluation provides a strong basis for evidence-based policy making, which is the modern standard in education governance. Policies are no longer based on assumptions or subjective reports, but on standardised measurements using valid and reliable instruments. For example, Aprianti's (2024) research shows that the process component in the implementation of Islamic Education courses in higher education is ineffective, so that study programmes can issue new policies related to teaching method innovation and digital technology integration.<sup>64</sup> Thus, Islamic education policies are more focused, measurable, and responsive to real needs in the field.

Furthermore, quantitative evaluation research enables the government to formulate clear performance indicators in Islamic education policy. Indicators such as teacher pedagogical competence, student religious literacy achievement, or curriculum effectiveness can be measured continuously and monitored over time. This process strengthens the control function of policy and ensures that national policy truly has an impact on improving the quality of Islamic education at the operational level. Empirical findings from various recent studies increasingly confirm that without strong quantitative data support, Islamic education policy risks being offtarget and difficult to evaluate systematically for effectiveness.

### **c. Implications for the Values and Philosophy of Islamic Education**

Quantitative evaluation research has profound implications for strengthening the values and philosophy of Islamic education, particularly in efforts to present an educational model that balances spiritual, intellectual, and social dimensions. The quantitative approach enables systematic measurement of abstract aspects of education, such as religious attitudes, values of moderation, and Islamic social behaviour. Through the use of valid and reliable psychometric instruments, these values can be evaluated objectively so that educational institutions gain a clear understanding of the level of internalisation of Islamic values among students.<sup>65</sup>

Furthermore, quantitative evaluation research reinforces the basic principles of Islamic educational philosophy, such as *tawazun* (balance) and *maslahah* (benefit). With empirical data, educational institutions can determine whether a programme truly provides tangible benefits for the spiritual and social development of students. For example, longitudinal research on the impact of Islamic education on students' religious attitudes and tolerance provides an overview of whether Islamic education is in line with the value of *al-wasathiyah* (moderation).<sup>66</sup> This type of evaluation ensures that Islamic education is not only normative but also proven to be effective in producing behavioural changes in line with Islamic pedagogical objectives.

From the epistemological perspective of Islamic education, the quantitative approach also enriches the scientific paradigm because it provides an objective mechanism for assessing the effectiveness of religious programmes, which have so far been evaluated more subjectively. This principle of objectivity does not conflict with Islamic values, but is in line with the concepts of *hisbah* and *muraqabah*, which

---

<sup>64</sup> N. Aprianti, "Evaluation Using the CIPP Model in Islamic Religious Education," ...h. 55–57.

<sup>65</sup> I. Tsani, "Affective Evaluation Model for Islamic Religious Education,"...h. 112–114.

<sup>66</sup> Khoridatul Bahiyah et al., "Learning Evaluation in Islam: Review of the Concept of the Qur'an and Hadith," *Modern Islamic Studies & Civilization* 5, no. 1 (2024)h. 88–92.

emphasise accountability and evidence-based assessment<sup>67</sup> Thus, the integration of quantitative evaluation research strengthens the philosophy of Islamic education, which emphasises harmony between normative vision (values) and empirical evidence (performance).

Ultimately, quantitative evaluation research reinforces the role of Islamic education as a system that is responsive to modern challenges without abandoning its religious identity. Data-based evaluation helps ensure that the idealism of Islamic education—such as the goal of forming people with noble character—does not stop at the conceptual level, but is reflected in measurable achievements that are scientifically and ethically accountable.<sup>68</sup>

#### **d. Implications for the Future Direction of Islamic Education Development**

The direction of Islamic education development in the future requires an evaluation approach that is not only descriptive but also capable of forming the basis for long-term strategic policy formulation. Quantitative evaluation research provides opportunities for Islamic educational institutions to implement evidence-based development, which is a decision-making process based on valid and measurable empirical data.<sup>69</sup> Thus, the direction of Islamic education development no longer relies solely on intuition, experience, or administrative tradition, but shifts towards data-driven governance.

The development of the PAI curriculum, for example, can be directed through the mapping of empirical competencies generated from standardised evaluation instruments. When the evaluation results show that certain competency achievements—such as religious literacy, understanding of religious moderation, or ethical thinking skills—are still low, then curriculum strengthening policies can be designed more accurately in these aspects<sup>70</sup> In this way, quantitative evaluation becomes a strategic instrument in ensuring that curriculum renewal is in line with the real needs of students.

In addition, the future direction of Islamic education development needs to strengthen the capacity of institutions and teachers in understanding evaluation data and using it as a basis for planning. Recent research shows that Islamic educational institutions that have a data-use culture tend to be more adaptive in responding to policy changes and learning needs.<sup>71</sup> Thus, strengthening data literacy for teachers, curriculum developers, madrasah principals, and management is an important asset in the development of Islamic education.

In a global context, Islamic education is also required to engage in the dynamics of digitalisation and technological transformation. Data-based evaluation enables educational institutions to monitor the effectiveness of the use of learning technologies, including PAI e-learning, Al-Qur'an learning applications, and religious

---

<sup>67</sup> Syed Muhammad Naquib al-Attas, *The Concept of Education in Islam* (Kuala Lumpur: ISTAC, 1991), h.45–46.

<sup>68</sup> M. Alamsyah, “Analisis Efektivitas Asesmen Kuantitatif dalam Meningkatkan Pendidikan Agama Islam,”... h. 45–47.

<sup>69</sup> Creswell, *Educational Research...*, h. 102–104.

<sup>70</sup> Daniel L. Stufflebeam and Anthony J. Coryn, *Evaluation Theory, Models, and Applications...* h. 36–38

<sup>71</sup> M. Alamsyah, “Analisis Efektivitas Asesmen Kuantitatif dalam Meningkatkan Pendidikan Agama Islam,... h. 46–47.

moderation platforms.<sup>72</sup> Findings from quantitative evaluation can be used to improve technology design, map digital challenges, and validate the most effective technology-based learning models for Muslim students.

Overall, quantitative evaluation research provides a new direction for the development of Islamic education that is more scientific, measurable, responsive, and needs-based. This new direction positions Islamic educational institutions as modern institutions capable of combining religious values with professional and accountable educational governance mechanisms.

#### **e. Final Synthesis**

The final synthesis of this study confirms that quantitative evaluation research has a strategic position in the development of Islamic education because it is able to provide an objective empirical basis for quality improvement, policy strengthening, and expansion of the direction of Islamic education reform. Through the integration of concepts, models, and empirical findings that have been described, it is evident that the quantitative approach functions not only as a measuring tool, but also as an epistemological device that supports the transformation of Islamic education towards a system that is more transparent, accountable, and adaptive to change. This approach enables Islamic educational institutions to identify quality gaps, assess the effectiveness of interventions, and formulate data-based policies that are accountable.

At the same time, this synthesis emphasises that quantitative evaluation research cannot stand alone without considering the socio-religious context and the basic values of Islamic education. Therefore, the integration of quantitative and qualitative approaches—as discussed in the previous section—becomes a more comprehensive evaluation framework that is in line with the holistic nature of Islamic education. From this perspective, quantitative evaluation provides objective measures, while qualitative evaluation provides a meaningful and contextual understanding. The combination of the two allows for the creation of policy recommendations that are not only scientific but also in line with the ethical and spiritual values of Islam.

Thus, the final synthesis of this study confirms that quantitative evaluation research serves as a methodological foundation for effective policy-making and development of Islamic education. Evaluation is no longer seen as merely administrative reporting, but as a scientific process that provides direction for change. This paves the way for Islamic education that is not only academically superior, but also relevant to the needs of the times, responsive to challenges, and remains adherent to the Islamic values that form its foundation.

### **E. Conclusion**

The results of this study indicate that quantitative evaluation research has a very strategic position in the development of Islamic education. Through an approach based on objective measurement and statistical analysis, quantitative evaluation is able to provide a measurable picture of the effectiveness of programmes, the quality of the learning process, and the achievement of Islamic education goals. Evaluation models such as CIPP, Kirkpatrick, Goal-Oriented, and Goal-Free provide a comprehensive framework that enables educational institutions to conduct assessments in a more systematic and targeted manner, literature findings

---

<sup>72</sup> N. Aprianti, "Evaluation Using the CIPP Model in Islamic Religious Education," *Conciencia* 13, no. 1 (2024): h. 55–56.

show that quantitative approaches are not only useful in evaluating cognitive dimensions, but can also be used to measure religious affective aspects through the development of valid and reliable psychometric instruments. On the other hand, empirical research in Islamic education shows that the use of quantitative evaluation can improve the quality of learning, strengthen programme accountability, and support data-based planning in madrasas, schools, universities, and non-formal educational institutions, however, quantitative evaluation research cannot stand alone.

The context of Islamic education, which is rich in values, meaning, and spiritual dimensions, requires a deeper qualitative understanding. Therefore, the complementary integration of quantitative and qualitative approaches is important to produce a comprehensive picture of programme effectiveness, both in terms of numerical achievements and process dynamics. This mixed approach ensures that the evaluation of Islamic education remains in line with modern scientific principles as well as ethical and religious values, overall, quantitative evaluation research makes an important contribution to improving the quality, accountability, and relevance of Islamic education. Through the application of appropriate evaluation models, standardised instruments, and evidence-based analysis, Islamic educational institutions can formulate more responsive policies, develop more effective programmes, and ensure the sustainability of quality education processes. This article is expected to serve as a conceptual basis for further research and a practical reference for policymakers in developing a more scientific, comprehensive, and sustainable evaluation of Islamic education.

## Bibliography

- Alamsyah, M. "Analisis Efektivitas Asesmen Kuantitatif dalam Meningkatkan Pendidikan Agama Islam." *Journal of Education Research and Policy* 4, no. 2 (2023): 45–47.
- Al-Attas, Syed Muhammad Naquib. *The Concept of Education in Islam*. Kuala Lumpur: ISTAC, 1991.
- Alanshori, M. Zainuddin, et al. "The Application of the CIPP Evaluation Model in Educational Programs in Indonesia." *Journal of Educational Evaluation* (2024): 6–7.
- Aprianti, N. "Evaluation Using the CIPP Model in Islamic Religious Education." *Conciencia* 13, no. 1 (2024): 55–57.
- Ary, Donald, et al. *Introduction to Research in Education*. 8th ed. Belmont: Wadsworth, 2010.
- Badan Litbang dan Diklat Kemenag RI. *Pedoman Asesmen Nasional Literasi Dasar Keagamaan (ANLDB)*. Jakarta: Puslitbang Pendidikan Agama dan Keagamaan, 2024.
- Bahiyah, Khoridatul, et al. "Learning Evaluation in Islam: Review of the Concept of the Qur'an and Hadith." *Modern Islamic Studies & Civilization* 5, no. 1 (2024): 88–92.
- Creswell, John W. *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research*. 6th ed. Boston: Pearson, 2023.
- \_\_\_\_\_. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. 5th ed. Thousand Oaks: SAGE Publications, 2018.
- Detik News. "Kemenag Gelar Asesmen Nasional Literasi Keagamaan 2025 bagi Guru dan Siswa SD." Diakses 16 Oktober 2025. <https://www.detik.com/hikmah/khazanah/d-8191892/kemenag-gelar-asesmen-nasional-literasi-keagamaan-2025-bagi-guru-dan-siswa-sd>.
- Direktorat Jenderal Pendidikan Islam. *Transformasi Digital Madrasah 2024: Arah Kebijakan dan Implementasi*. Kementerian Agama RI. Diakses 11 Oktober 2025. <https://madrasah.kemenag.go.id/berita/transformasi-digital-madrasah>.
- Guba, Egon G., dan Daniel L. Stufflebeam. *Evaluation in Education: The Need and the Response*. Bloomington: Phi Delta Kappa, 1970.
- Guskey, Thomas. "Data-Driven Decision Making and the Future of Educational Evaluation." *Educational Measurement: Issues and Practice* 42, no. 1 (2023): 12–19.
- Kementerian Agama Republik Indonesia. "Moderasi Beragama: Kunci Indonesia Emas 2045 – Penjelasan Direktur Diktis." Diakses 10 Oktober 2025. <https://pendis.kemenag.go.id/direktorat-perguruan-tinggi-keagamaan-islam/moderasi-beragama-kunci-indonesia-emas-2045-ini-penjelasan-direktur-diktis>.
- Kementerian Agama RI. *Moderasi Beragama: Buku Saku Penguatan Moderasi Beragama*. Jakarta: Kemenag RI, 2022.

- \_\_\_\_\_. *Roadmap Transformasi Digital Madrasah 2023–2029*. Jakarta: Direktorat KSKK Madrasah, 2023.
- Kirkpatrick, Donald L., dan James D. Kirkpatrick. *Evaluating Training Programs: The Four Levels*. 3rd ed. San Francisco: Berrett-Koehler, 2006.
- Miles, Matthew B., A. Michael Huberman, dan Johnny Saldaña. *Qualitative Data Analysis*. 4th ed. Thousand Oaks: SAGE Publications, 2020.
- Moore, Gary W. *Developing and Evaluating Educational Research*. Boston: Little, Brown and Company, 1983.
- Muhammad Ali. *Memahami Riset Perilaku dan Sosial*. Bandung: Pustaka Cendekia Utama, 2011.
- Mukhlis, M. “The Effectiveness of the Lok-R Model in Enhancing Academic Achievement in the Islamic Religious Education Study Program.” *AL GHAZALI: Jurnal Pendidikan Dan Pemikiran Islam*. 2025.
- Mukhlis, Mukhlis, Ahyar Rasyidi, and Husna Husna.. “Tujuan Pendidikan Islam: Dunia, Akhirat Dan Pembentukan Karakter Muslim Dalam Membentuk Individu Yang Berakhlak Dan Berkontribusi Positif.” *AL GHAZALI: Jurnal Pendidikan Dan Pemikiran Islam* 1–20. doi: 10.69900/ag.v4i1.189. 2024
- Rossi, Peter H., Howard E. Freeman, and Mark W. Lipsey. *Evaluation: A Systematic Approach*. 7th ed. Thousand Oaks: SAGE Publications, 2004.
- Scriven, Michael. *Evaluation Thesaurus*. 4th ed. Newbury Park: SAGE Publications, 1991.
- Stufflebeam, Daniel L., dan Anthony J. Coryn. *Evaluation Theory, Models, and Applications*. 2nd ed. San Francisco: Jossey-Bass, 2014.
- Tashakkori, Abbas, dan Charles Teddlie. *Mixed Methodology: Combining Qualitative and Quantitative Approaches*. Thousand Oaks: SAGE Publications, 1998.
- Tsani, I. “Affective Evaluation Model for Islamic Religious Education.” *Didaktika* 12, no. 2 (2023): 112–114.
- Tyler, Ralph W. *Basic Principles of Curriculum and Instruction*. Chicago: University of Chicago Press, 1949.
- Zuhdi, Muhammad. “Evaluasi Kebijakan Pendidikan Islam: Tantangan dan Prospek.” *Jurnal Pendidikan Islam* 13, no. 2 (2023): 221–240.