

# KORESPONDENSI

komentar reviewer 1

## DEVELOPING DELTA INTERNAL QUALITY ASSURANCE TO EVALUATE QUALITY OF INDONESIAN ISLAMIC UNIVERSITY

SitiChoiriyah,Kumaidi&BadrunKartowagiran  
State University of YogyakartaYogyakarta, Indonesia  
Corresponding author: [choiriain@gmail.com](mailto:choiriain@gmail.com)

**Abstract--adjusted to IJET styles, it is Emerald.**

**Purpose:** This study report: 1) aspects of internal quality assurance to evaluate the Indonesian Islamic universities, 2) the development of Delta internal quality assurance (DIQA) to achieve a standard model of evaluation, and 3) evidences to developed DIQA appropriate to evaluate the quality of Indonesian Islamic universities.

**Design/Methods/Approach:** This study is a Research and Development (R&D) in context CIPP (context, input, process, product and product). Delphi method was used to collect data. This study applied four cycles: exploration, preliminary testing, main field testing field-testing, and operational-main operational testing. The development of Development-process was done through the Delphi method to obtain the model and statistical analysis to validate the items in the model.

**Findings:** This study revealed the prototype of the DIQA model instrument has been named as DIQA, the DIQA model is improved improved based on the main field and operational main-testings. The DIQA is completed with statistical analysis to describe the validity and reliability of each item. The final Final version of DIQA has seven7 dimensions of evaluation, ten10 kinds of questionnaire and 477 items of questions.

**Implication:** Devised to accommodate accommodate Islamic values and provided a high degree of internal quality, DIQA entails challenges challenge to align with national accreditation system by the government. It implies that DIQA should receive receive more dissemination dissemination factually and publicly, encouraging Islamic university to use DIQA confidently.

**Originality/Originality/Values:** DIQA has its own specialty in promoting Islamic values in the evaluation of internal quality assurance. The accuracy and propriety propriety of the instrument has shown showed specific dimension that another other instrument may not yet cover.

**Keywords:** internal quality, benchmarking, quality assurance assurance, Islamic values.

### 1. INTRODUCTION

This article is an extract of doctoral dissertation reporting the development of an instrument to evaluate internal quality assurance for the Islamic university in Indonesia. The development focuses on how the instrument meets the needs of internal quality assurance of study programs, evaluating the vision and mission, curriculum, teaching learning process, infrastructure, facilities and students' outcomes. The instrument is developed using the Delphi Delphi method, so we label it as the Delta Internal Quality Assurance (henceforth known as DIQA). This study is conducted for three reasons: quality quality assurance is a pivotal concern in higher education management (Jelena & Hećimović, 2016; Tam, 2001); quality assurance is the way benchmarking is launched (Shafer & Coate, 1992); and, DIQA as a tool to evaluate internal quality assurance for the Islamic university is convincing (Choiriyah, 2018).

Formatted: Left: 3 cm, Right: 3 cm, Top: 3 cm, Bottom: 3 cm, Width: 21 cm, Height: 29,7 cm

Formatted: Centered

Formatted: Font color: Custom Color(RGB(255;0;102))

Formatted: Font: Bold, Font color: Custom Color(RGB(255;0;102)), Complex Script Font: Bold

The issue of the quality of higher education and the role of higher education institutions in quality assurance and continuous improvement has been formally adopted within the framework of education development policy (Jelena & Hećimović, 2016, p. 75). The key for quality assurance is to inform the widest possible academic community, starting from teachers to students and the administrative staff that will put quality assurance into practice and implement it in their respective institution (Tam, 2001).

In the global expansion, higher education has put information on academic quality as a benchmarking for the university. Universities must introduce systematic evaluations of education at departmental, faculty and university-wide levels (Rossi et al., 2004). Higher education, like industry, has to pose benchmarking practices in its operations with specific performance targets. The benchmarking process assists to identify and understand the drivers of processes, outputs and quality. It provides objective measurements for goal setting to enable a university track or find the extent of meeting the set targets (Shafer & Coate, 1992). Benchmarking provides managers of an institution with an external point of reference or standard for evaluating the quality and cost of their organization's internal activities, practices, and processes (Tam, 2001).

Attempts to measure quality in higher education should be based on the purpose and major goals. The assessment programme should reflect what constitutes the quality in higher education, determines the outcomes to be measured, and the approach of measuring them (Tam, 2001). Priorities of quality focus on three basic goals of higher education: research, public or community service, and education of students. The education of students should be the primary objective, which gives reasons for the existence of universities. Students are a major part of the concept and universities are required to provide quality education by making optimal favourable conditions to promote effective learning. Hence, for any consideration of quality, the improvement of the student experience should be of central importance (Patil & Pudlowski, 2005, 2005; Tam, 2001).

Indonesia is now facing problems, challenges and issues concerning low quality of input, process to output quality, and outcome of school graduates, besides, universities have not been able to fulfil the industrial needs because they are late to respond to the development of information and technology (Fitri, 2016, p. 206). Many higher education institutions are not accredited, and there is an acute shortage of advanced human capital. Accreditation capacity must be strengthened, and stronger regulation is needed to address low-quality providers. To achieve the world benchmarks, it will be necessary to improve investment and internationalisation of research capacity among universities (OECD, 2015, p. 20).

Accreditation has been made, but shortages appear, addressing that the Board of National Accreditation for Higher Education (BAN-PT) is not consistent enough to apply the quality assurance. The process of accreditation is time-consuming, disconnecting of the monitoring process, product oriented and focusing more on the administrative that does not focus on the quality assurance (Fitri, 2016). Thus, it is necessary for the Islamic university to devise their own instruments that allow training for stakeholders to build educational quality culture. This study, therefore, is

intended to develop DIQA, an instrument to evaluate the internal quality assurance for the Islamic university. Specifically, three research ~~questions~~ questions guide the investigation.

- 1) What aspects of internal quality assurance are determined to evaluate Indonesian Islamic universities to operate standard services of a university teaching-learning programs?
- 2) What are the development processes of DIQA to achieve a standard model of evaluation?
- 3) What evidences are developed to strengthen the Delta internal quality assurance appropriate to assess the internal quality of Indonesian Islamic universities?

## 2. Review of Literature

### 2.1. Internal Quality Assurance

To conceive internal quality assurance, the quality is firstly defined. Various ways of defining quality have evolved in the literature. Watty (2006, p. 293) noted quality is about efficiency, high standards, excellence, value and value for money, fitness and fitness for purpose and/or customer focused. Quality is fitness for purpose that includes mission, goals, objectives, specifications and specifications. Fitness for purpose means that an organization has procedures that are appropriate for the specified purposes, and that the procedures are achieving the specified purposes. Quality ~~is~~ has four components as defined by Harvey and Green (1993) in Table 1.

**Table 1: Harvey and Green's (1993) classification of quality**

Classification	Brief explanation
Quality as exceptional	A focus on meeting high standards, such as excellence
Quality as perfection or consistency	As embodied in the idea that something is done correctly or to a consistent standard every time
Quality as fitness for purpose	Where quality is defined in terms of the achievement of a desired educational or quality assurance goal
Quality as value for money	A focus on ensuring that stakeholders receive high value for their investment
Quality as transformation	A focus on ensuring that students are genuinely empowered as a result of their learning

In higher education institutions, quality assurance has been made as a mechanism to control quality. Essentially, higher education ~~s~~ undertakes s major reforms in ~~their~~ structures and activities as an impact of globalization, accountability, supply and demand issues, competition, and technology. The maintenance, improvement, and assurance of quality for higher institutions have become a major concern and has have come to the attention of ~~to~~ governments and other stakeholders. Higher Education Evaluation Council (HEEC, 2008, p. 9) ~~admits~~ admits. "quality Quality assurance means the procedures, processes, and systems used by the higher education institution to manage and improve the quality of its education and other activities." Quality

assurance should confirm that the higher education has adequate conditions or provisions in place to enable students to achieve the set standards.

Quality assurance has been described as: "All attitudes, objects, actions, and procedures which together with the quality control activities, ensure that appropriate academic standards are being maintained and enhanced in and by the ~~program~~programme, institution or system, and make this known to the educational community and the public at large." (Woodhouse, 1999, p. 30). Approaches in quality assurance vary ~~from accreditation~~from accreditation, assessment, academic audit and external examination. Each practice allows the development and setting of the criteria and the application of those criteria or set standards to a programme or institution by the accrediting body. The purpose may be assessment or enhancement with the aim of further improvement of the programme or the educational system at large (Lenn, 2004).

Arcaro (1995, p. 1) suggests a quality program basically includes ~~four~~ ~~components~~four components: commitment to change, understand well the condition of the program or institution, have a clear vision of the future and everyone in the institution must stick to that vision, and have plans to implement the quality of educational institutions. BAN-PT (2015) classifies quality educational institutions as: (1) shared understanding and commitment to high goals, (2) open communication and collaborative problem solving, (3) continuous assessment for teaching and learning, (4) personal and professional learning, (5) resources to support teaching and learning, and (6) curriculum and instruction.

## 2.2. Practices of Quality Assurance

Practices of quality assurance relate to ~~the~~ assessment and benchmarking. Competitive pressure to achieve universal access makes the assessment of higher education institutions a major concern for the public (Tam, 2001; Patil & Pudlowski, 2012). Koslowski (2006) suggests that like in industries, the higher education views quality measurable product or service and is achieved when expectations or requirements are met. Quality represents products, service, and knowledge that ~~is~~ ~~are~~ ~~basically~~ ~~evaluated~~basically evaluated by customer satisfaction. To Koslowski, university quality is ~~determined~~ ~~by~~ ~~determined~~ by its outputs, such as efficient use of resources and whether or not it produces competent, highly satisfied and employable graduates. ~~Quality is defined by the customer~~The customer defines quality; management is responsible for the quality, and how quality can be improved.

Koslowski (2006) asserts the quality of the process is when the higher education institutions view the work as valuable, measurable, and improvable. Assessment is a ~~measurable~~ ~~process~~measurable process that aims ~~to improve~~improving quality; an assessment is parts of evaluating quality. ~~Assessment in higher education~~ ~~are~~Assessment in higher education is guided self-assessment, intermediary conduct assessment, independent self-assessment, and student competencies-~~based~~ ~~assessment~~based assessment. Guided ~~self-assessment~~self-assessment is based on peer review similar to a business certification such as the International Standards Organization: ISO 9000. Koslowski (2006) believes that the academic audit has become a dominant model for institutional assessment in higher education. Through the independent self-assessment, higher education institutions assess the needs of customers, the process of education and results.

Internal quality assurance (Utuka, 2012; Tam, 2001) includes a quality assurance policy that is publicly available and part of strategic planning (Tam, 2001); design and

Formatted: Font color: Auto

Formatted: Font color: Blue

approval of programs suggesting ~~that the~~ that the program design allows to meet the set objectives and the intended learning outcomes, such as regulation of student admission, progression, recognition and certification (Tam, 2001; Patil&Pudlowski, 2005); teaching: teaching staff, learning resources and student support; effective management of their programs, publish information about their activities, including programs, which should be clear, accurate, objective, up-to date and readily accessible (Tam, 2001); and monitoring and periodically review of the programs to ensure that they are achieving the set objectives and respond to the needs of students and society (Utuka, 2012;Patil&Pudlowski, 2005, Tam, 2001).

Internal quality assurance in universities (BAN-PT, 2015) can be controlled through various models of quality management, with PDCA model (Plan, Do, Check, Action) that results in continuous improvement or high quality kaizen, being frequently used. PDCA-based quality control management works on the following principles.

- 1). Quality first: All thoughts and actions of education managers should be prioritized on quality.
- 2). All for stakeholders: All thoughts and actions of education managers must have a purpose to give satisfaction to stakeholders.
- 3). Our stakeholders: Any person performing duties in any process undertaken by higher education should consider others who will use their work as stakeholders to be satisfied.
- 4). Speak with data: Any action and decision taken in the processes at the college should be based on analyzing the data that has been collected and processed, not based on supposition or engineering.
- 5). Upstream management: All decision-making in the higher education process is done in a participatory, not authoritative way.

Indonesian HELTS (Higher Education Long Terms Strategy) 2003-2010 Article 2 statesstates. “~~In a~~ In a ~~healthy~~ healthy organization, ~~a~~ a ~~continuous~~ continuous quality improvement should become ~~its~~ its ~~primary~~ primary concern. Quality assurance should be internally driven, institutionalized within each organization’s standard procedure, and involve external parties. However, since quality is also a concern of all stakeholders, quality improvement should aim at producing outputs and outcomes as part of public accountability” (BAN-PT, 2015). Quality assurance in higher education systems comprises of internal and external quality assurance systems. However, the implementation of internal quality assurance by BAN-PT is elaimed ~~not~~ claimed not fully ~~in~~ practice ~~in~~ practice compared to the external quality assurance activities (Haris, 2013).

In 2008, according to ~~Harris~~ Haris (2013), BAN-PT has done ~~the~~ a first national assessment on the implementation of internal quality assurance at some of universities in Indonesia. This focus is more on the existence and completeness of the institutional documents regarding internal quality assurance activities in each university or academic institution. Basically, three major objectives are done in the internal quality assurance, they are:

- 1) Assessing ~~the~~ to what extent ~~extend~~ of the implementation of internal quality assurance in higher education.
- 2) Presenting a quality profile of every unit of learning in the university, to expose the strength and weakness of its quality assurance program.

**Commented [1]:** A non-English word must be defined in parantheses the first time you use it in a text. Please revise.

**Formatted:** Font: Not Italic, Font color: Auto, Complex Script Font: Not Italic

3) Recording feedback, suggestions and recommendations to the universities that have implemented the internal quality assurance to improve, develop and straighten the implementation of internal quality assurance in their institution.

In the accreditation process, Kelchen (2017) suggests accreditors typically judge a college based on five broad standards.

- The college’s mission must be appropriate for the accreditor.
- The college must have adequate governance structures and an independent governing board.
- The college must demonstrate financial health—the ability to continue operating throughout the accreditation cycle. This is the most common reason colleges are at risk of losing recognition (GAO 2014).
- The college must have sufficient academic resources, including faculty members, facilities, and library resources.
- In the 1980s accreditors ~~used~~ ~~use~~ ~~student learning~~ ~~student-learning~~ outcomes as a standard, because explicit standards were not set, the implications of this change are unclear (Ewell 2010).

Formatted: Font color: Auto

## 2.2. Quality Assurance and Measurement

Quality assurance and quality measurement was used when the growth of higher education began and the structure of the higher education sector became more complex (Tam, 2001; Keltchen, 2017). The internationalisation process of higher education, and the introduction of free trade economy has made education providers to place the quality in the world market, emphasizing the standard activity of the input, process and output for the quality assessment of education structures (Anderson, Johnson, and Milligan, 2000; Patil&Pudlowski, 2005, p. 52). Any standard of industrial activity includes three different stages, such as the input, the process and the output (Anderson, Johnson, and Milligan, 2000). In this process, feedback gained from the output can be utilised to improve the quality of the process. This model has also been adopted for the quality assessment of education structures. The three stages of an educational process cycle are further elaborated by Anderson, Johnson, and Milligan (2000) in Figure 1.

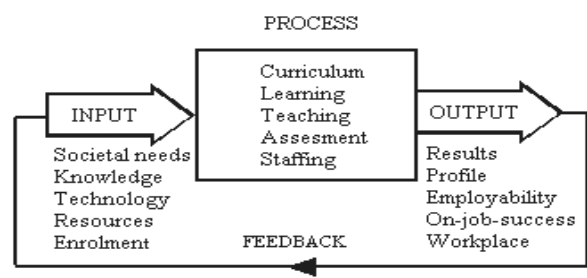


Figure 1: The block diagram of an educational cycle (adopted from Patil&Pudlowski, 2004 & Anderson, Johnson, and Milligan, 2000).

Patil&Pudlowski (2004) and Anderson, Johnson, and Milligan (2000) elaborate educational input, process and output as follows.

### Educational Input

The *Input* parameters relate to student's intake or student's enrolment into an educational process, comprising the following aspects:

- Societal needs;
- New knowledge;
- Advancing technologies;
- Human and material resources;
- Student enrolment process;
- Student fees structure;
- Student eligibility criteria

Formatted: Justified, After: 0,03 cm, Space Before: 0 pt

### Educational Process

The educational process lies in between the input and the output, and this is where teaching/learning is facilitated. It may consist of the following important factors:

- Curriculum design;
- Learning styles;
- Learning methods;
- Teaching/learning facilities;
- Assessment methods;
- Staffing.

Formatted: Font: Not Italic, Complex Script Font: Not Italic

### Learning Outcomes

The *Output* component is associated with the student output after finishing the course curricula. It consists of the following elements:

- Academic results;
- Professional profile;
- Employability;
- On-the-job success rate;
- Social and workplace activities, etc.

### 3. ~~Measurement~~ Measurements and Benchmarking in Quality Assurance

Measurement and benchmarking are two inseparable brands in the quality assurance. "What you measure is what you get" (Chinta and Jelena, 2016, p.989). "What benchmarks you use is what meaning you get" (Chinta&Jelena, 2016, p. 990). Measurement becomes the basis for utilizing multiple metrics in performance management to ensure ~~that organizations~~ that organizations seek to achieve progress along multiple dimensions (Podsakoff et al., 2000; Chinta and Jelena, 2016). In addition, benchmarking is ~~evaluating~~ evaluating ~~an action~~ an action with a ~~standard~~ standard ~~for~~ standard for comparison. Employee engagement is enhanced with greater shared understanding of the metrics used (Rich et al., 2010; Chinta and Jelena, 2016).

There are factors affecting the quality of the education system regarding benchmarking. In general, the factors are: effective learning and teaching, leadership, lecturers, students, institutional management, physical environment and resources, stakeholder satisfaction, institutional culture, learning outcomes or performance, and accountability (Bridge, Judd, and Mook, 1979: 1-3). Drawn in more comprehensive context, there are ten indicators including: (1) effective learning and teaching (20%), (2) leadership (15%), (3) Staff (15%), (4) students (15%), (5) standards (10%), (6) organization (5%), (7) physical environment and resources (5%), (8) external



relations (5%), 9) access (5%), and (10) service to customers (5%) (Sallis (2002: 151-159). Finally, BAN-PT (2015) defines the quality is identified from six indicators: (1) shared understanding and commitment to high goals, (2) open communication and collaborative problem solving, (3) continuous assessment for teaching and learning, (4) personal and professional learning, (5) resources to support teaching and learning, and (6) curriculum and instruction.

Higher education institutions compete globally because of student mobility on an international scale. Universities concentrate on using available resources and quality assessment using continuous quality improvement (CQI) and people-oriented (Shakhnoza, 2009; Roffe, 1998). Roffe (1998, p.74) defines that CQI comes from Japanese term *kaizen* which *kaizen*, which means “slow never-ending improvements in all aspects of life”. He distinguishes traditional CQI from the modern one. The traditional CQI is the famous classic Western approach: spending large amounts of money on improving the quality and purchasing new technology. In the Japanese context CQI, or *kaizen*, is continuously making small steps to improve the existing system and equipment by people who manage or work in the system. The structural steps to this CQI or *kaizen* are as follows:

- defining the area of improvement
- analyzing and selecting appropriate problems
- identifying causes
- planning counter-measures
- implementing
- confirming the results
- standardization (Roffe, 1998, p.75).

### 3. Methods

#### 3.1. Approach

This study is a Research and Development (R&D) that uses a qualitative and quantitative approach. The study adapts the R&D approach from Borg and Gall (1983; 2003). The steps are: (1) Research and information collecting, (2) Planning, (3) Develop a preliminary form of the product; (4) Preliminary field testing, (5) Main product revision; (6) Main field testing; (7) Operational product revision; (8) Operational field testing; (9) Final product revision, and (10) ~~Dissemination~~ ~~Dissemination~~ and ~~implementation~~ ~~implementation~~. In the development process to achieve the prototype, we used a two round Delphi method. In addition, CIPP (context, input, process, product) is implemented as the framework of evaluation study (Rita & Shokrpour, 2011). Basically, this study ~~applied~~ ~~applied~~ three research stages: the exploration stage to develop prototype, the trials of the prototype to improve into a model and revision of the model. The research was conducted in the Faculty of Islamic Education and Teaching (FITK) of Islamic State Institute (IAIN) Surakarta from January to November 2017. As the instrument was developed under the procedures of Delphi method, we called the product as Delphi Internal Quality Assurance (henceforth is abbreviated as DIQA).

#### 3.2. Participants

This study involved 222 participants, consisting of academic experts, management of six study programs, lecturers, and students. The participants were recruited using purposive sampling techniques and they were involved in three ~~stages~~: exploration, preliminary testing, main field testing, and operational main testing. Table 2 describes the participants.

Commented [2]: Please move this to the first place you use the Japanese term.



Table 2. Participants and their roles in the research stage

No	Research stages	Participants				Total
		Experts	Lecturer	Management	Students	
1	Exploration	2	5	5	10	22
2	Preliminary testing, Delphi method	10	10	10	NA	30
3	Main field testing	-	10	10	50	70
4	Operational main testing	-	10	10	100	120
						222

### 3.3. Research Procedures

Procedures of research indicated data collection in four stages of the research. In the exploration stages, data ~~was were collectedeollected~~ through observation and interview with experts, lecturers, management and students in a ~~restrictedrestruced~~ format. The aim was to ~~identifyidentify~~ needs, weaknesses and expectation of how internal assurance was formatted. This stage revealed needs analysis, prototype of instrument for internal quality assurance and criteria of evaluation.

In the preliminary testing, prototype of instrument was evaluated and improved through ~~the~~ Delphi method ~~inthat used~~ two rounds. ~~The DelphiThe methodDelphi ismethod —anis —establishedan —researchestablished —methodologyresearch aimedmethodology —specificallyaimed —atspecifically at~~ exploring the expected future of novel and evolutionary phenomena. The technique obtains a group of experts' most reliable consensus of opinion (Sekayi& Kennedy, 2017) by allowing them to express their own views on a topic, while taking into account the other participants' views by means of controlled feedback. The method is based on the premise that well-informed individuals, drawing on their insights and on prior experience, are better equipped to predict the future than theoretical approaches or extrapolation of trends (Grobelaar, 2007). The responses to ~~thea series ofquestionnaires arequestionnaires —anonymousare anonymous.Participants areParticipants —alsoare —providedalso —withprovided —awith summarya —ofsummary of~~ opinions from a previous round before answering the next questionnaire. It is believed that such a consensus process will converge the group toward the 'best' response.

In the first round, the prototype was submitted to 30 members participating in the Delphi. Along with the submission, a questionnaire to assess the quality of the prototype was attached. The members also provided comments that identified the weakness of the prototype and the ways it was improved. All comments and suggestions from the experts were used as the main data to improve the prototype. In the second round, members of ~~the~~ Delphi method were invited ~~to meet together~~ to discuss the results obtained in the first round. The results were a ~~consensuseonceneus~~ that after the improvement, the prototype was feasible to define ~~as~~ the model of internal quality assurance for Islamic university.

In the main field ~~testing~~ and operational ~~main~~ testing, data ~~was were~~ focused on the results of DIQA using ~~statisticalstatisteal~~ analysis using SPSS to provide evidences of validity, ~~reliabilityreliability~~, conformity of each item of the DIQA, analyzing the construct using Confirmatory Factor Analysis (CFA), and hypothesis testing model. Test of suitability of model, validity and reliability of this construct are statistically

analyzed using CFA with the help of LISREL 8.70 program (Ghozali and Fuad, 2005, pp. 29-34). The test include:

- 1) Goodness of Fit Index (GFI), an Index that describes the overall suitability of the model compared to the actual data. The GFI value > 0.90 suggests good suitability.
- 2) Root Mean Square Error of Approximation (RMSEA) to indicate the fit model with value < 0.05. RMSE 0.08 to 1.0 sufficient fit.
- 3) Normed Fit Index (NFI), a comparative measure between the proposed model and the null model. The recommended value is NFI > 0.90.

Techniques of data analysis in this section is summarized in Table 3.

Table 3. Summary of Techniques of data analysis

Techniques of analysis	Application
Descriptive statistics using	Calculating mean, percentage and determining criteria obtained from expert validation, Delphi technique, instrument sheet scores from reviewer
Using SPSS 17 and CFA using Lisrel program	Construct validity testing instrument of DIQA obtained from main field testing and operational main testing
Construct Alpha using SPSS 17	Reliability testing of the instrument obtained from main field testing and operational main testing.

#### 4. Results and Discussion

This study contributes three findings: the results of the development of the prototype of DIQA, trials of the DIQA, and statistical analysis to achieve the final product of DIQA.

##### 4.1. The Development of the DIQA Prototype

This longitudinal research was conducted from January to October 2017. In general, the research procedures cover four main actions: exploration and developing the initial draft of the prototype, preliminary testing using the Delphi method, main field testing, and operational field testing. The instrument was developed under the Delphi method so we named it the Delta Internal Quality Assurance (DIQA). CIPP (context, input, process, product) was used as the model of evaluation in the DIQA.

Preparation of the development of the prototype of the instrument of internal quality assurance for Islamic higher education began from the exploration process from which we developed the needs analysis. The exploration assessed the available quality instruments available and confirmed through interview to 6 staff members and 8 leaders at IAIN Surakarta. Documents pertaining to curriculum, quality assurance, staffs, students, and implementation of overall management assurance in 6 study programs were evaluated.

The results approved needs analysis and an initial draft to develop the prototype of DIQA. The needs analysis covers seven components that CIPP evaluation should cover: (1) Vision and mission of the program, (2) Curriculum, (3) Competency of lecturers and administration staff, (4) Infrastructure and facilities, (5) Teaching learning process, (6) Students atmosphere supervision, and (7) Graduate learning outcomes. Preliminary, we developed 10 kinds of questionnaires and 477 final items, criteria of evaluation, and scoring techniques. This way, a two-round Delphi

Formatted: Indent: Before: 0 cm, First line: 0 cm

Formatted: Font color: Auto

Formatted: Font color: Auto

Formatted: Font color: Auto

Formatted: Font color: Auto

Formatted: Font color: Auto

technique was used. The result of this stage agreed to receive the draft as a prototype of DIQA.

Coverage of the prototype appears in Table 4. Results of the preliminary test using the Delphi method are summarized in table 5 for general appearance and Table 6 to evaluate the content coverage. Criteria to evaluate the results appear in Table 7.

**Table 4. Evaluation object and number of items in DIQA.**

No	Evaluation object	Number of items
1	Vision and mission document	7
2	Curriculum document	12
3	Lecturer competency	85
4	Competency of staff and administration	32
5	Infrastructure and Facilities	184
6	Document of teaching learning planning	15
7	Teaching learning process	26
8	Assessment of teaching and learning	12
9	Document of students supervision	30
10	Graduate competency	77
Total		<b>480</b>

**Table 5. General Appearance of DIQA**

Indicator	ore	%	Criteria
appearance	33	94	Very good
if evaluation	32	91	Very good
description of the components	30	86	Very good
ility	32	91	Very good
understand	30	86	Very good
systematic	34	97	Very good
ge use	31	89	Very good
of writing	32	91	Very good
election, font, and spacing	33	4,3	Very good
ss of pages	30	5,7	Very good
ility do answer	29	83	Very good
fectiveness to do	30	86	Very good
ion achievement	32	91	Very good

Of 24 aspects, seven reaches a percentage above 76% (very good), and for aspects: page thickness, evaluation guide, time to work, implementation evaluation, data analysis, criteria determination, and evaluation report reporting above 51%. In terms of practicality, DIQA is categorized as practical (85% ease and 89% benefit). DIQA is categorized as efficient (86%) and deeper than AMI (89.5%). Of the seven internal quality standards of DIQA are divided into six books, namely: (1) Book 1: questionnaire document evaluation of mission vision, curriculum, lesson planning and student coaching document; (2) Book 2: questionnaire of lecturer competence; (3) Book 3: questionnaire of employee competency evaluation; (4) Book 4: questionnaire evaluation of facilities; (5) Book 5: questionnaire of process evaluation and assessment of learning, and (6) Book 6: questionnaire of graduate competency evaluation. Separation into six books aims to make the evaluation more efficient if it is based on the evaluation objective and reduces the "bold" impression on previous packaging evaluation models.

Formatted: Font color: Auto

Formatted: Indent: First line: 0 cm

Formatted: Font: Bold, Complex Script Font: Bold

Formatted: Font: Bold, Complex Script Font: Bold

Formatted: Font color: Auto

Formatted: Font color: Auto

Formatted: Font color: Auto

**Table 6. Evaluation of Instrument Dimension**

Component	Dimension	Score	%	Criteria
A. Vision and mission of study of study program	1. Vision and Mission of Islamic Study program	33	94,3	Very good
B. Curriculum	1. Curriculum of Islamic education	32	91	Very good
C. Competency of lecturers and Administration staff	1. Lecturer competency	27	90	Very good
	2. Competency of Administration staff	32	91	Very good
D. Infrastructure and facilities	1. General standard	34	97	Very good
	2. Mosque	30	86	Very good
	3. Classroom	29	83	Very good
	4. Library	31	89	Very good
	5. Laboratory	24	79	Good
	6. Leader room	25	80	Good
	7. Lecturer room	34	97	Very good
	8. Administration room	33	94	Very good
	9. Toilet	33	94	Very good
E. Teaching learning Process	1. Teaching-learning plan	33	94	Very good
	2. Teaching implementation	34	97	Very good
	3. Assessment	29	83	Very good
F. Students guidance	1. Aims and objectives	32	91	Very good
	2. Kinds of students guidance	29	97	Very good
G. Graduate competency	1. Personality competency	28	93	Very good
	2. Pedagogic competency	32	91,4	Very good
	3. Professional competency	29	96,7	Very good
	4. Social competency	27	90	Very good

**Table 7. Degree of component in DIQA**

Percentage	Criteria
1% - 20,99%	Very poor
21% - 40,99%	Poor
41% - 60,99%	Moderate
61% - 80,99%	Good
81%-100%	Very good

## 4.2. Main Field Testing and Operational Main Testing

### 4.2.1. Descriptive analysis of DIQA

The descriptive analysis is used to report quantitatively the result of rates for the DIQA. Result of analysis appears in Table 8.

Formatted: Font color: Auto

Table 8. Descriptive analysis of instrument in Main Field Testing

Parts to be evaluated	Mean	%	Criteria
and mission of study program	3,826	96,5	Very good
structure and facilities	3,826	96,5	Very good
teacher competency	3,315	82,88	Very good
administration staff competency	3,681	92,03	Very good
curriculum	3,826	96,5	Very good
staff supervision	3,826	96,5	Very good
content of teaching learning process	3,826	96,5	Very good
method	3,826	96,5	Very good
teaching learning process	3,536	88,40	Very good
graduate competency	3,674	91,50	Very good

Formatted: Indent: First line: 0 cm

Evidence shows all instruments reached >76%—that meaning means the DIQA is "very good". Substantially, improvements are required as suggested by respondents as follows: (1) Answers to questions should be made in options a, b, c, and d, so that the respondents can easily cross-mark option they deem fit, (2) For factual data, respondents are expected to provide brief description to meet factual evaluation data; (3) Self assessment for students should include options students can select, (4) Indicators of research and devotion of pedagogical competencies are inserted into professional competence; (5) Three indicators of learning planning are made into two only; (6) Comfort and space security indicators are added for the items of infrastructure and facilities.

Table 9. Result of quality evaluation of all study program programs in FITK IAIN Surakarta

Evaluation	No	Name	Mean	Category
Input	1	Vision and mission	3,83	Very good
	2	Competency of <del>lecturer</del> <del>and lecturer and</del> administration staff	3,31	Very good
	3	Curriculum	3,66	Very good
	4	Infrastructure and <del>Facilities</del> Facilities	3,54	Good
<b>Mean of Input</b>			<b>3,36</b>	Very good
Process	5	Teaching learning process	3,73	Very good
	6	Supervisory	3,58	Very good
	<b>Mean of Process</b>			<b>3,65</b>
Output	7	Competency of graduate	3,34	Good
<b>Mean of output</b>			<b>3,34</b>	Good
<b>Mean of Evaluation</b>			<b>3,45</b>	Very good

Formatted: Font: 10 pt, Bold, Complex Script Font: 10 pt, Bold

Formatted: Font: 10 pt, Bold, Complex Script Font: 10 pt, Bold

In summary, the average result of the quality evaluation of the study program in FITK is "excellent" with a score of 3.45. Between input, process, and output evaluation in the DIQA model, there has a comprehensive linkage, proving that good output is also determined by excellent input and process quality.

Of 24 aspects evaluated by DIQA, 20 reached above 76% (excellent), and five that include page thickness, evaluation guidance, data analysis, criteria determination, and preparation of evaluation report reached above 51% (good). Refer to table 9.

Table 10

Result of review of all instruments for PBA, PBI, PGMI and TBI

GENERAL FORMAT				
Indicator	Max	Score	%	Criteria
Design & appearance of the model	54	57	105,94	Very attractive
Clarity of writing	54	51	94,27	Very good
Consistency of words, font, and spacing	54	48	88,9	Very good
Systematic	54	48	88,9	Very good
Page use	50	44	88	Very good
Thickness of page	54	47	87,0	Fairly thick
Readability	54	52	96,3	Easy to read
Easy to understand	54	47	87,0	Easy to understand
SUBSTANCE OF THE MODEL				
Instructional guide	50	49	98	Easy to understand
Clarity of evaluation	50	46	92	Clarity of evaluation already covered
Clarity of component description	54	44	81,5	Component described
Clarity to do the instrument	54	52	96,3	Easy to understand
Work	54	47	87,0	Easy to do
Effectiveness of work	50	41	82	Time effectiveness
Usefulness	54	43	79,6	Very useful
Clarity of evaluation	54	40	74,1	Used to evaluate school
Clarity of evaluation instrument	54	43	79,6	Used to evaluate study program
Comparison to internal quality audit (AMI)	44	35	79,5	Easy to use
Comparison to other evaluation model	48	33	68,8	Easier to use
PROCEDURE OF EVALUATION				
Clarity and planning	52	48	92,3	Practical
Clarity of evaluation	56	53	94,6	Very easy to do
Clarity of evaluation data	52	49	94,2	Easy to do
Clarity of evaluation criteria	52	47	90,4	Easy to do
Clarity to result of evaluation	48	45	93,8	Easy to do

In general, this means that DIQA shows that it is (92%) and efficient (86.7%) as compared with AMI. Further test is required to see the validity and reliability empirically.

#### 4.2.2. Results of Confirmatory Factor Analysis

Extensive operational trials were conducted in four study programs: PBA (Arabic Study Program), PBI (English Language Education Study Program), PGMI (Teacher Education for Elementary Islamic School), and TBI (Unit of Indonesian Language Program), Faculty of Education and Teaching IAIN Surakarta involving 242 respondents.

Formatted: Indent: First line: 0 cm

#### 1) CFA Test to Vision and Mission Aspect

Vision and mission section consists of the vision and mission of the study program and its goals. Results of CFA on mission and vision with 7 items are satisfied. The measurement model achieves good fit: CFI = 0.97 and value > 0.9 which means

goodness of fit well. In addition, t values for all items are greater than 1.00, which means that they are generally compatible with the mission aspect aspect theory. This indicates that the 7 points are valid statement points for construct measurement from aspect of vision and mission of study program.

## 2) CFA Curriculum Aspects

The curriculum aspect consists of curriculum design and curriculum criteria, measured by 12 items. Results of CFA indicate the value of CFI = 1.00 and a value > 0.9 which means DIQA model has a good of fit. The t values of all items greater than 1.96 means all items in general conform with curriculum construct in the items.

## 3) CFA aspects of Competencies of Lecturers and Administration Staff

~~Competency of lecturer and administration staff is measured by 7 points~~ Seven points measure competency of lecturer and administration staff. Value of the good fit is  $p = 0.31018$  ( $p > 0.05$ ) ~~which means these items~~, which means these items, are good to measure. In addition, the t values for all items, that are greater than 1.96, means that the items are generally compatible with the theory competency of lecturer and employee. Evidently, ~~the,~~ the items are valid for the construct measurement of lecturer and staff competencies.

## 4) CFA for Infrastructure and Facilities

Measurement of CFA for infrastructure and facilities is measured by nine points, namely general sarpras, mosque, space, library, laboratory, leadership room, faculty room, administration room, and toilet. Results of test using GFI = 0.94, AGFI = 0.93, NFI = 0.94 and CFI = 1.00 are value > 0.9. This means the items have the goodness of fit. In addition, the results of t values for all items are greater than 1.96, which means all items are generally in conformity with the infrastructure stated in the DIQA model; the items contain valid statement points for the construct measurement of infrastructure aspects.

**Commented [3]:** This phrase is not used in standard English. Please consider rewriting for clarity.

## 5) CFA Model Quality Evaluation Process Aspects of Learning

The learning process has three components namely, planning, implementation and assessment and it contains 12 items. Respectively, results of test using GFI = 0.94, AGFI = 0.93, NFI = 0.94 and CFI = 1.00 are value > 0.9, meaning the items have the goodness of fit. In addition, t-values for all items greater than 1.96 means that the items are generally compatible with the ethical and regulatory aspects of theory. The items are valid statement points for construct measurement of aspects of the learning process.

## 6) CFA Aspects of Student Development

Aspects of student coaching consists of guidance, guardianship, skill practice, literary reading and bilingual are measured with 30 items of statement. CFA results indicating GFI = 0.97, AGFI = 0.97, NFI = 0.93 and CFI = 0.97 have a value > 0.9. This means the items in DIQA have a goodness of fit. The t-values for all grains that are value greater than 1.96, means the grains are in conformity with the theory of student coaching aspects. The items are valid statement points for the construct measurement of student coaching aspects.



## 7) CFA Aspects of Graduate Competence

The CFA measurement for the graduate competency consists of 4 components, namely personality competency, pedagogical ~~competence~~competency, professional ~~competence~~competency and social competency component and is measured by 15 items. The results CFA are GFI = 0.97, AGFI = 0.97, NFI = 0.93 and CFI = 0.97 with a value > 0.9. It means the items have the goodness of fit. In addition, the t values for all grains that are greater than 1.96, means the grains are generally in conformity with the theory of student coaching aspects. The items are valid statement points for the construct measurement of student coaching aspects.

### 4.2.3. Result of ~~Hypothetical~~Hypotetic Model Testing DIQA

The modelling ~~hypothesis~~hypothetis testing of DIQA shows evidence that DIQA necessitates all components of input, process, and output evaluation to be evaluated. Evaluation of input quality will determine the quality of process, and quality of the process affects the output.

**Input Quality Evaluation.** This evaluation is to examine the quality of Islamic education study programs by looking at: (a) vision and mission of the study program; (b) the curriculum and curriculum design; (c) competence of lecturers and employees related to pedagogical competence, professional competence, social competence and personality competence; and (c) infrastructure and facilities, such as mosques, classrooms, libraries, ~~multimedia~~multi-media, laboratories, leadership rooms, faculty rooms, administration rooms and toilets.

**Quality Process Evaluation.** This evaluation is to assess the quality of the study program by looking at: (a) implementation of learning related to the planning, ~~the~~ implementation process of learning and assessment; (b) student coaching related to thesis guidance, guardian of study, Al-Qur'an literacy coaching, expertise practice, and language development of students.

**Evaluate Output Quality.** This evaluation is to see the quality of graduates by measuring teacher professionalism that includes pedagogical competence, professional competence, social competence and personality competence.

The statistical analysis results are clarified below. The DIQA test with CFA using SEM, proves that DIQA has good ability in matching data (good fit). Evidences on the standardized loading of hypothetical model of component relations, ~~variables~~ of input quality evaluation, process quality and output quality, show that correlational indicators among variables have a high loading factor  $\geq 0.3$  (Tabachnick&Fidell, 2007: 217; Donna, 2009: 215). This means that the main indicator of latent construct of DIQA model has been ~~well-rated~~well-rated and understood by the respondents. The DIQA model constructs are ~~well-applied~~well-applied and highly deserve to be maintained and used.

The loading factor value means (1) quality evaluation of input to process quality has a loading factor value of 0.32 with the quadratic value of  $0.32^2 = 10.24$ . This means that a 10.24% variant of input quality influences process quality, and (2) evaluation of process quality to output quality has a loading value equal to 0.57 with ~~the~~ quadratic value ( $0.57^2 = 0.3249$ ). Also, 32,49% variant of process quality evaluation influences output quality. Thus, evaluation of quality of input affects the quality of process and ultimately contributes quality evaluation of teacher quality output. This result is reinforced with t-value with cut-off 5% (value  $t = 1.96$ ). The evidence implies (1)

quality of input significantly influences process quality, (2) evaluation of process quality significantly influences output quality.

### 5. Conclusion and Implication

In summary, this study contributes three findings: kind of needs for the internal quality assurance, process of development of DIQA and empirical evidence to achieve final product of DIQA model.

- 1) Needs analysis is useful for internal quality assurance in the Islamic university ~~including include~~: vision and mission of study program, curriculum, competency for lecturer and administration staff, infrastructure and facilities, students coaching, teaching learning process, assessment for teaching learning process, and graduate competency. DIQA has ~~ten+0~~ different sets of questionnaire and total items 480. DIQA is based on the CIPP to develop the constructs, method of evaluation, and procedures to include research procedures.
- 2) The development of items is begun from the exploration stage to prepare the needs analysis and initial package of the DIQA. Validation of items starts from the preliminary testing through the Delphi method conducted in two rounds. ~~The results of Results~~ the Delphi method are used to improve the prototype into a model. Revision addresses that DIQA has ~~seven+7~~ dimension dimensions of evaluation, ~~ten+0~~ kinds of questionnaires, and 477 items as a strong format. Inspired by the Delphi method, the model is named ~~as DIQA, that~~ DIQA that stands for Delta Internal Quality Assurance.
- 3) DIQA has served to improve for qualitative and quantitative techniques. Qualitative approach has been used to develop the prototype during preliminary testing and main field-testing to improve the model of DIQA. Quantitatively, statistical analysis using SPSS and ISREL is prepared to see evidence on validity, reliability, items and options in each questionnaire. The results evidently prove that the appearance of DIQA is very good, individual items are valid, and consistency to evaluate the study program is reliable. The final version of DIQA approves to consist of seven dimensions of internal quality assurance, ten kinds questionnaire, and 477 items, improving questionnaire from AMI and BAN-PT.

This research, however, ~~experienced receives~~ limitations: less cooperative respondents, uncertain timing of the evaluation in the study program, dilemma to determine respondents for measurement of the competence graduates, and positioning DIQA with government accreditation. Less cooperative respondents have made the objectives of evaluation are not clear and not fully match. Time of evaluation frequently does not conform so that the external ~~validity validity~~ may shed. Determining student graduates who actually represent graduate competence is frequently more subjective. Tracer ~~studies study~~ that serve comprehensive data on graduate competency ~~are is~~ not fully ~~available available~~. Finally, DIQA cannot relocate government evaluation for accreditation.

The limitations ~~give arise~~ implications for the study program and future research. First, benchmarking is the ultimate goal of accreditation. The benchmarking through accreditation makes the study program ~~to~~ strive to get a good accreditation value meeting all indicators required in accreditation. Thus, the Islamic education program must follow the rules or standards set by BAN PT, though a specificity and peculiarity of Islamic values are not covered.

Formatted: Font color: Auto

Formatted: Font color: Auto

Formatted: Font color: Auto

Formatted: Font color: Auto

Formatted: Font color: Auto

Formatted: Font color: Auto

Formatted: Font color: Auto

Formatted: Font color: Auto

It implies that the study programs can make use of DIQA to accommodate Islamic values, as the internal quality assurance instrument. DIQA serves the internal quality evaluation for the Islamic education program in producing professional teachers and helps prepare the external evaluations, such as accreditation. DIQA is evidently an effective way to know the quality of Islamic education programs in producing professional teachers because it is able to: 1) produce a study profile with up to date data and information; 2) planning and improving the quality of Islamic education programs on a regular basis; 3) provide information about the community and parties in need; and 4) prepare the study program in the face of external evaluation or accreditation.

The second implication that arises pertaining to the future research. This study has obtained evidence that feeling reluctant of the reluctance of respondents to be involved in the research and uncertain time to conduct the evaluation affect the validity of the items. This implies that DIQA contains less comprehensive attributes of accreditation and aligning DIQA with BAN-PT is something problematic. Future research can verify the items of DIQA improving the items and dimension of evaluation. Efforts to align DIQA as an the initial accreditation training are recommended for use by the Islamic universities.

Formatted: Font color: Auto

Formatted: Font color: Auto

#### **REFERENCES** ----will be revised based on the clean version after final edit.

Abadie-Mendia, Trudy, "Accountability and Quality in Higher Education: A Case Study " (2013). *UNF Theses and Dissertations*. 375. University of North Florida. <http://digitalcommons.unf.edu/etd/375>

Arcaro, S.J. (1995). *Quality in Education: An Implementation Handbook*. Alih bahasa Yosallriantara. Yogyakarta: Pustaka Pelajar

BAN-PT (National Accreditation Body for Higher Education). (2002). *Guidelines for Internal Quality Assessment of Higher Education*. Second edition. Ministry of National Education.

Bridge, R.G. Judd, C.M. & Mook, P.R. (1979). *The Determinants of Educational Outcomes*. Massachusetts: Ballinger Publishing Company

Chinta, Ravi; Kebritchi, Mansureh & Ellias, Janelle. (2016). A conceptual framework for evaluating higher education institutions. *International Journal of Educational Management*, Vol. 30 No. 6, pp. 989-1002. DOI: 10.1108/IJEM-09-2015-0120.

Choiriyah, Siti. (2018). *Model Evaluasi Mutu Internal Program Studi Pendidikan Islam dalam Menghasilkan Guru Profesional*. Disertasi Doktor Tidak Diterbitkan. Yogyakarta: Program Studi Penelitian dan Evaluasi Pendidikan, Universitas Negeri Yogyakarta.

Directorate General of Higher Education. (2003). Higher Education Long Term Strategies: *HELTS 2003 - 2010*. Department of National Education, Jakarta.

Donna, H. (2009). *Confirmatory factor analysis*. Oxford: Oxford University Press.

- Ewell, Peter. 2010. "Twenty Years of Quality Assurance in Higher Education: What's Happened and What's Different?" *Quality in Higher Education* 16 (2): 173–75.
- Fitri, ZainulAgus. (2016). Quality Assurance System between the Islamic University and the State University. *Aljamiah*. Vol. 2 No. 2. pp. 208-235. DOI: <http://dx.doi.org/10.15575/jpi.v2i2.787>.
- GAO (US Government Accountability Office). (2014) and Andrea Fuller and Douglas Belkin, "The Watchdogs of College Education Rarely Bite," *Wall Street Journal*, June 17, 2015, <https://www.wsj.com/articles/the-watchdogs-of-college-education-rarely-bite-1434594602>.
- Gaspersz, V. (2005). *Total Quality Management*. Jakarta: PT GramediaPustakaUtama
- Ghozali, I. (2009). *Aplikasi Analisis Multivariate dengan Proses SPSS*. Semarang: Badan Penerbit Universitas Diponegoro
- Grobelaar, S. (2007). *Data Gathering: Delphi Method*. R&D in the National System of Innovation: A System, Dynamic Model. University of Pretoria Etd. Pretoria, South Africa.
- Haris, Ikhfan. (2013). Assessment on the Implementation of Internal Quality Assurance at Higher Education (An Indonesian Report). *Journal of Educational and Instructional Studies in the World*. Vol. 3 No 4, pp. 45-49.
- Harvey, L. and Green, D. (1993) Defining quality. *Assessment and Evaluation in Higher Education*, 18 (1): 9-34.
- Kelchen, Robert. (2017). *Higher Education Accreditation and Federal Government*. Washington DC, Urban Institute. Urban.org.
- Koslowski III, F.A. (2006). Quality and assessment in context. *Quality Assurance in Education*, 3, 277-288. doi:10.1108/09684880610678586
- Lenn, M. P. (2004). *Quality assurance and accreditation in higher education in East Asia and the Pacific*. Washington, DC: The World Bank.
- MSCHE (Middle States Commission on Higher Education). 2014. "Standards for Accreditation and Requirements of Affiliation." 13th ed. Philadelphia: MSCHE.
- NEA (National Education Association). (2006). *School quality: Issues in education*. (<http://www.nea.org/jobs/>). Diambil dari <http://www.proquest.com/pqdweb.html>
- Patil, S. Arun & Pudlowski, J. Zenon. (2005). Important Issues of the Accreditation and Quality Assurance and a Strategy in the Development of Accreditation Framework for Engineering Course. *Global Journal of Engineering Education*. Vol. 9 No. 1, pp. 49-62.

- Podsakoff, P.M., MacKenzie, S.B., Paine, J.B. and Bachrach, D.G. (2000), "Organizational citizenship behaviors: a critical review of the theoretical and empirical literature and suggestions for future research", *Journal of Management*, Vol. 26 No. 3, pp. 513-563.
- Rich, B.L., Lepine, J.A. and Crawford, E.R. (2010), "Job engagement: antecedents and effects on job performance", *Academy of Management Journal*, Vol. 53 No. 3, pp. 617-635.
- Rita, Rezaee & Shokrpour, Nasrin. (2011). Performance Assessment of Academic Departments: CIPP Model. *European Journal of Social Sciences*, Vol. 23 No. 2. pp. 228-236.
- Roffe, I. M. (1998). Conceptual problems of continuous quality improvement and innovation in higher education. *Quality Assurance in Education*, 6, 74-82.
- Rossi, P. H., Lipsey, M. W., & Freeman, H. E. (2004). *Evaluation: A systematic approach* (7th ed<sup>7th</sup>ed.). CA: Sage Publications
- Sallis, E. (2002). *Total quality management in education*. London: Kogan Page Limited
- Sekayi, D., & Kennedy, A. (2017). Qualitative Delphi Method: A Four Round Process with a Worked Example. *The Qualitative Report*, 22(10), 2755-2763. Retrieved from <http://nsuworks.nova.edu/tqr/vol22/iss10/15>
- Shafer, B. S., & Coate, L. E. (1992). Benchmarking in higher education: A tool for improving quality and reducing cost. *Business Officer*, 26(5), 28-35.
- Shakhnoza (2009). *Perception Of Quality In Changing University Education In Kazakhstan*. Yakubova: Kent State University College and Graduate School of Education, Health, and Human Services
- Tabachnick, B. G., & Fidell, L. S. (2007). *Experimental designs using ANOVA*. Colombia: Thomson Brooks/Thomson & Brooks/Cole
- Tam, Siu-ling, Maureen. (2001). *Assessing Quality in Higher Education by Examining the Effect of University Experiences on Learning Outcomes and Student Development*. PhD<sup>PhD</sup>s Dissertation. University of Durham. Stockton, Durham City, UK.
- Utuka, Godwin. (2012). *Quality Assurance in Higher Education: Comparative Analysis of Provision and Practices in Ghana and New Zealand*. PhD<sup>PhD</sup>s Dissertation. Wellington: Victoria University of Wellington.

Van Vught F.A., & Westerheijden D.F. (1994). Towards a general model of quality assessment in higher education. *Higher Education*, 28, 355-371.

Watty, K. (2006). Want to know about quality in higher education? Ask an academic. *Quality in Higher Education*, 12(3), 291-301.

Woodhouse, D. (1999). Quality and quality assurance: an overview. In H. De Wit & J. Knight (Eds.), *Quality and internationalisation in higher education* (pp. 29-43). Paris: OECD.

komentar reviewer 2

## DEVELOPING DELTA INTERNAL QUALITY ASSURANCE TO EVALUATE QUALITY OF INDONESIAN ISLAMIC UNIVERSITY

SitiChoiriyah, Kumaidi&BadrunKartowagiran  
State University of Yogyakarta, Indonesia  
Corresponding author: [choiriain@gmail.com](mailto:choiriain@gmail.com)

**Abstract---adjusted to IJET styles, it is Emerald.**

**Purpose:** This study report: 1) aspects of internal quality assurance to evaluate the Indonesian Islamic universities, 2) the development of Delta internal quality assurance (DIQA) to achieve a standard model of evaluation, and 3) evidence to develop DIQA appropriate to evaluate the quality of Indonesian Islamic universities.

**Design/Methods/Approach:** This study is a Research and Development (R&D) in context CIPP (context, input, process, and product). Delphi method was used to collect data. This study applied four cycles: exploration, preliminary testing, main field-testing, and main operational testing. The development process was done through the Delphi method to obtain the model and statistical analysis to validate the items in the model.

**Findings:** This study revealed the prototype of the DIQA model is improved based on the main field and operational [testingstesting](#). The DIQA is completed with statistical analysis to describe the validity and reliability of each item. The final version of DIQA has seven dimensions of evaluation, ten kinds of questionnaire and 477 items of questions.

**Implication:** Devised to accommodate Islamic values and provide a high degree of internal quality, DIQA entails challenges to align with national accreditation system by the government. It implies that DIQA should receive more dissemination factually and publicly, encouraging Islamic university to use DIQA confidently.

**Originality/Values:** DIQA has its own specialty in promoting Islamic values in the evaluation of internal quality assurance. The accuracy and propriety of the instrument has shown specific dimension that another instrument may not yet cover.

**Keywords:** internal quality, benchmarking, quality assurance, Islamic values.

### 1. INTRODUCTION

This article is an extract of a doctoral dissertation reporting the development of an instrument to evaluate internal quality assurance for an Islamic university in Indonesia. The development focuses on how the instrument meets the needs of internal quality assurance of study programs, [evaluating-evaluates](#) the vision and mission, curriculum, teaching-learning process, infrastructure, facilities and students' outcomes. The instrument is developed using the Delphi method, so we label it as the Delta Internal Quality Assurance (henceforth known as DIQA). This study is conducted for three reasons: quality assurance is a pivotal concern in higher education management (Jelena&Hećimović, 2016; Tam, 2001); quality assurance is the way benchmarking is launched (Shafer &Coate, 1992); and, DIQA as a tool to evaluate internal quality assurance for the Islamic [university-University](#) is convincing (Choiriyah, 2018).

The issue of the quality of higher education and the role of higher education institutions in quality assurance and continuous improvement has been formally adopted within the framework of education development policy (Jelena&Hećimović, 2016, p. 75). The key for quality assurance is to inform the widest possible academic community,

Formatted: Indent: Before: 0 cm

Formatted: Left: 3 cm, Right: 3 cm, Top: 3 cm,  
Bottom: 3 cm, Width: 21 cm, Height: 29,7 cm



starting from teachers to students and the administrative staff that will put quality assurance into practice and implement it in their respective institution (Tam, 2001).

In the global expansion, higher education has put information on academic quality as a benchmark for the university. Universities must introduce systematic evaluations of education at departmental, faculty and university-wide levels (Rossi et al. 2004). Higher education, like industry, has to pose benchmarking practices in its operations with specific performance targets. The benchmarking process assists to identify and understand the drivers of processes, outputs, and quality. It provides objective measurements for goal setting to enable a university track or find the extent of meeting the set targets (Shafer & Coate, 1992). Benchmarking provides managers of an institution with an external point of reference or standard for evaluating the quality and cost of their organization's internal activities, practices, and processes (Tam, 2001).

Attempts to measure quality in higher education should be based on the purpose and major goals. The assessment programme should reflect what constitutes the quality of higher education, determines the outcomes to be measured, and the approach of measuring them (Tam, 2001). Priorities of quality focus on three basic goals of higher education: research, public or community service, and education of students. The education of students should be the primary objective, which gives reasons for the existence of universities. Students are a major part of the concept and universities are required to provide quality education by making optimal favourable conditions to promote effective learning. Hence, for any considerations of quality, the improvement of the student experience should be of central importance (Patil & Pudlowski, 2005; Tam, 2001).

Indonesia is now facing problems, challenges and issues concerning low quality of input, process to output quality, and outcome of school graduates, besides, universities have not been able to fulfil the industrial needs because they are late to respond to the development of information and technology (Fitri, 2016, p. 206). Many higher education institutions are not accredited, and there is an acute shortage of advanced human capital. Accreditation capacity must be strengthened, and stronger regulation is needed to address low-quality providers. To achieve the world benchmarks, it will be necessary to improve investment and internationalisation of research capacity among universities (OECD, 2015, p. 20).

Accreditation has been made, but shortages appear, addressing that the Board of National Accreditation for Higher Education (BAN-PT) is not consistent enough to apply the quality assurance. The process of accreditation is time-consuming, disconnecting of the monitoring process, product-oriented and focusing more on the administrative that does not focus on the quality assurance (Fitri, 2016). Thus, it is necessary for the Islamic university to devise their own instruments that allow training for stakeholders to build educational quality culture. This study, therefore, is intended to develop DIQA, an instrument to evaluate the internal quality assurance for the Islamic University. Specifically, three research questions guide the investigation.

- 1) What aspects of internal quality assurance are determined to evaluate Indonesian Islamic universities to operate standard services of a university teaching-learning programs?
- 2) What are the development processes of DIQA to achieve a standard model of evaluation?

3) What evidences are developed to strengthen the Delta internal quality assurance appropriate to assess the internal quality of Indonesian Islamic universities?

**2. Review of Literature-----ok agree to revise adhering your consideration. Thank you.**

**2.1. Internal Quality Assurance**

To conceive internal quality assurance, the quality is ~~firstly~~, defined first. Various ways of defining quality have evolved in the literature. Watty (2006, p. 293) noted quality is about efficiency, high standards, excellence, and value for money, and fitness for purpose and/or customer focused. Quality is fitness for purpose that includes mission, goals, objectives, and specifications. Fitness for purpose means that an organization has procedures that are appropriate for the specified purposes and that the procedures are achieving the specified purposes. Quality has four components as defined by Harvey and Green (1993) in Table 1.

**Commented [ST1]:** Some of the material in 2.1 and 2.2 is quite repetitive, I would consider making cuts or rewording here to shorten the overall length of the piece.

**Table 1: Harvey and Green’s (1993) classification of quality**

<b>Classification</b>	<b>Brief explanation</b>
Quality as exceptional	A focus on meeting high standards, such as excellence
Quality as perfection or consistency	As embodied in the idea that something is done correctly or to a consistent standard every time
Quality as fitness for purpose	Where quality is defined in terms of the achievement of a desired educational or quality assurance goal
Quality as value for money	A focus on ensuring that stakeholders receive high value for their investment
Quality as transformation	A focus on ensuring that students are genuinely empowered as a result of their learning

In higher education institutions, quality assurance has been made as a mechanism to control quality. Essentially, higher education undertakes major reforms in structures and activities as an impact of globalization, accountability, supply and demand issues, competition, and technology. The maintenance, improvement, and assurance of quality for higher institutions have become a major concern and have come to the attention of governments and other stakeholders. Higher Education Evaluation Council (HEEC, 2008, p. 9) admits, “Quality assurance means the procedures, processes, and systems used by the higher education institution to manage and improve the quality of its education and other activities.” Quality assurance should confirm that the higher education has adequate conditions or provisions in place to enable students to achieve the set standards.

Quality assurance has been described as: “All attitudes, objects, actions, and procedures which together with the quality control activities, ensure that appropriate academic standards are being maintained and enhanced in and by the program, institution or system, and make this known to the educational community and the public at large.” (Woodhouse, 1999, p. 30). Approaches in quality assurance vary from accreditation, assessment, academic audit and external examination. Each practice

allows the development and setting of the criteria and the application of those criteria or set standards to a programme or institution by the accrediting body. The purpose may be assessment or enhancement with the aim of further improvement of the programme or the educational system at large (Lenn, 2004).

Arcaro (1995, p. 1) suggests a quality program ~~basically~~ includes four components: commitment to change, understand well the condition of the program or institution, have a clear vision of the future and everyone in the institution must stick to that vision, and have plans to implement the quality of educational institutions. BAN-PT (2015) classifies quality educational institutions as: (1) shared understanding and commitment to high goals, (2) open communication and collaborative problem solving, (3) continuous assessment for teaching and learning, (4) personal and professional learning, (5) resources to support teaching and learning, and (6) curriculum and instruction.

## 2.2. Practices of Quality Assurance

Practices of quality assurance ~~relate-related~~ to assessment and benchmarking. Competitive pressure to achieve universal access makes the assessment of higher education institutions a major concern for the public (Tam, 2001; Patil & Pudlowski, 2012). Koslowski (2006) suggests that like in industries, ~~the~~ higher education views quality measurable products or services and is achieved when expectations or requirements are met. Quality represents products, service, and knowledge that are basically evaluated by customer satisfaction. To Koslowski, university quality is determined by its outputs, such as efficient use of resources and whether or not it produces competent, highly satisfied and employable graduates. The customer defines quality; management is responsible for the quality, and how quality can be improved.

Koslowski (2006) asserts the quality of the process is when the higher education institutions view the work as valuable, measurable, and improvable. Assessment is a measurable process that aims to improve quality; an assessment is a part of evaluating quality. Assessment in higher education is guided self-assessment, intermediary conduct assessment, independent self-assessment, and student competencies-based assessment. Guided self-assessment is based on peer review similar to a business certification such as the International Standards Organization: ISO 9000. Koslowski (2006) believes that the academic audit has become a dominant model for institutional assessment in higher education. Through the independent self-assessment, higher education institutions assess the needs of customers, the process of education and results.

Internal quality assurance (Utuka, 2012; Tam, 2001) includes a quality assurance policy that is publicly available and part of strategic planning (Tam, 2001); design and approval of programs suggesting that the program design allows to meet the set objectives and the intended learning outcomes, such as regulation of student admission, progression, recognition and certification (Tam, 2001; Patil & Pudlowski, 2005); teaching staff, learning resources and student support; effective management of their programs, publish information about their activities, including programs, which should be clear, accurate, objective, up-to date and readily accessible (Tam, 2001); and monitoring and periodically review of the programs to ensure that they are achieving the set objectives and respond to the needs of students and society (Utuka, 2012; Patil & Pudlowski, 2005, Tam, 2001).

Internal quality assurance in universities (BAN-PT, 2015) can be controlled through various models of quality management, with PDCA model (Plan, Do, Check,

Action) that results in continuous improvement or high-quality ~~kaizen~~. OK AGREE.  
PDCA-based quality control management works on the following principles.

**Commented [S12]:** A non-English word must be defined in parentheses the first time you use it in a text. Please revise.

- 1). Quality first: All thoughts and actions of education managers should be prioritized on quality.
- 2). All for stakeholders: All thoughts and actions of education managers must have a purpose ~~to give of giving~~ satisfaction to stakeholders.
- 3). Our stakeholders: Any person performing duties in any process undertaken by higher education should consider others who will use their work as stakeholders to be satisfied.
- 4). Speak with data: Any action and decision taken in the processes at the college should be based on analysing the data that has been collected and processed, not based on supposition or engineering.
- 5). Upstream management: All decision-making in the higher education process is done in a participatory, not authoritative way.

Indonesian HELTS (Higher Education Long-Term Strategy) 2003-2010 Article 2 states, “In a healthy organization, a continuous quality improvement should become its primary concern. Quality assurance should be internally driven, institutionalized within each organization’s standard procedure, and involve external parties. However, since quality is also a concern of all stakeholders, quality improvement should aim at producing outputs and outcomes as part of public accountability” (BAN-PT, 2015). Quality assurance in higher education systems comprises of internal and external quality assurance systems. However, the implementation of internal quality assurance by BAN-PT is claimed not fully in practice compared to the external quality assurance activities (Haris, 2013).

In 2008, according to Harris (2013), BAN-PT ~~had~~ done the first national assessment on the implementation of internal quality assurance at some of ~~the~~ universities in Indonesia. This focus is more on the existence and completeness of the institutional documents regarding internal quality assurance activities in each university or academic institution. Basically, three major objectives are done in the internal quality assurance, they are:

- 1) Assessing the extent of the implementation of internal quality assurance in higher education.
- 2) Presenting a quality profile of every unit of learning in the university, to expose the strength and weakness of its quality assurance program.
- 3) Recording feedback, suggestions, and recommendations to the universities that have implemented the internal quality assurance to improve, develop and straighten the implementation of internal quality assurance in their institution.

In the accreditation process, Kelchen (2017) suggests accreditors typically judge a college based on five broad standards.

- The college’s mission must be appropriate for the accreditor.
- The college must have adequate governance structures and an independent governing board.
- The college must demonstrate financial health—the ability to continue operating throughout the accreditation cycle. This is the most common reason colleges are at risk of losing recognition (GAO 2014).
- The college must have sufficient academic resources, including faculty members, facilities, and library resources.

- In the 1980s accreditors used student-learning outcomes as a standard, because explicit standards were not set, the implications of this change are unclear (Ewell 2010).

## 2.2. Quality Assurance and Measurement

Quality assurance and quality measurement ~~were~~ used when the growth of higher education began, and the structure of the higher education sector became more complex (Tam, 2001; Keltchen, 2017). The internationalisation process of higher education, and the introduction of free trade economy has made education providers to place the quality in the world market, emphasizing the standard activity of the input, process and output for the quality assessment of education structures (Anderson, Johnson, and Milligan, 2000; Patil & Pudlowski, 2005, p. 52). Any standard of industrial activity includes three different stages, such as the input, the process and the output (Anderson, Johnson, and Milligan, 2000). In this process, feedback gained from the output can be utilised to improve the quality of the process. This model has also been adopted for the quality assessment of education structures. The three stages of an educational process cycle are further elaborated by Anderson, Johnson, and Milligan (2000) in Figure 1.

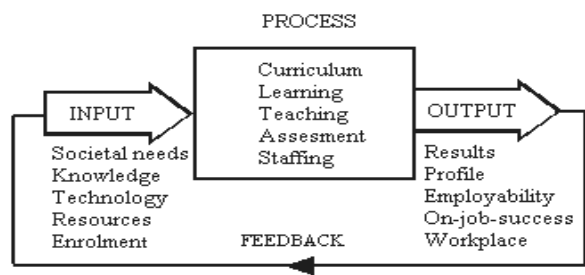


Figure 1: The block diagram of an educational cycle (~~adapted~~ adopted from Patil & Pudlowski, 2004 & Anderson, Johnson, and Milligan, 2000).

Patil & Pudlowski (2004) and Anderson, Johnson, and Milligan (2000) elaborate educational input, process, and output as follows.

### Educational Input

The *Input* parameters relate to student's intake or student's enrolment into an educational process, comprising the following aspects:

- Societal needs;
- New knowledge;
- Advancing technologies;
- Human and material resources;
- Student enrolment process;
- Student fees structure;
- Student eligibility criteria

### Educational Process

The educational process lies in between the input and the output, and this is where teaching/learning is facilitated. It may consist of the following important factors:

- Curriculum design;
- Learning styles;
- Learning methods;
- Teaching/learning facilities;
- Assessment methods;
- Staffing.

### Learning Outcomes

The *Output* component is associated with the student output after finishing the course curricula. It consists of the following elements:

- Academic results;
- Professional profile;
- Employability;
- On-the-job success rate;
- Social and workplace activities, etc.

### 3. Measurements and Benchmarking in Quality Assurance

Measurement and benchmarking are two inseparable brands in the quality assurance. “What you measure is what you get” (Chinta and Jelena, 2016, p.989). “What benchmarks you use is what meaning you get” (Chinta&Jelena, 2016, p. 990). Measurement becomes the basis for utilizing multiple metrics in performance management to ensure that organizations seek to achieve progress along multiple dimensions (Podsakoff et al., 2000; Chinta and Jelena, 2016). In addition, benchmarking is evaluating an action with a standard for comparison. Employee engagement is enhanced with a greater shared understanding of the metrics used (Rich et al., 2010; Chinta and Jelena, 2016).

~~There are factors affecting the quality of the education system regarding benchmarking. In general, the~~ The factors affecting the quality of the education system regarding benchmarking are: effective learning and teaching, leadership, lecturers, students, institutional management, physical environment and resources, stakeholder satisfaction, institutional culture, learning outcomes or performance, and accountability (Bridge, Judd, and Moock, 1979: 1-3). Drawn in more comprehensive context, there are ten indicators including: (1) effective learning and teaching (20%), (2) leadership (15%), (3) Staff (15%), (4) students (15%), (5) standards (10%), (6) organization (5%), (7) physical environment and resources (5%), (8) external relations (5%), 9) access (5%), and (10) service to customers (5%) (Sallis (2002: 151-159). Finally, BAN-PT (2015) defines the quality is identified from six indicators: (1) shared understanding and commitment to high goals, (2) open communication and collaborative problem solving, (3) continuous assessment for teaching and learning, (4) personal and professional learning, (5) resources to support teaching and learning, and (6) curriculum and instruction.

Higher education institutions compete globally because of student mobility on an international scale. Universities concentrate on using available resources and quality assessment using continuous quality improvement (CQI) and people-oriented (Shakhnoza, 2009; Roffe, 1998). Roffe (1998, p.74) defines that CQI comes from Japanese term *kaizen*, which means “slow never-ending improvements in all aspects of life.” He

**Commented [ST3]:** Please move this explanation to the first place you use the Japanese term.

distinguishes traditional CQI from the modern one. The traditional CQI is the famous classic Western approach: spending large amounts of money on improving the quality and purchasing new technology. In the Japanese context CQI, or *kaizen*, is continuously making small steps to improve the existing system and equipment by people who manage or work in the system. The structural steps to this CQI or *kaizen* are as follows:

- defining the area of improvement
- analyzing and selecting appropriate problems
- identifying causes
- planning counter-measures
- implementing
- confirming the results
- standardization (Roffe, 1998, p.75).

### 3. Methods

#### 3.1. Approach

This study is a Research and Development (R&D) that uses a qualitative and quantitative approach. The study ~~adopts~~<sup>adapts</sup> the R&D approach from Borg and Gall (1983; 2003). The steps are: (1) Research and information collecting, (2) Planning, (3) Develop a preliminary form of the product; (4) Preliminary field testing, (5) Main product revision; (6) Main field testing; (7) Operational product revision; (8) Operational field-testing; (9) Final product revision, and (10) Dissemination and implementation. In the development process to achieve the prototype, we used a two-round Delphi method. In addition, CIPP (context, input, process, product) is implemented as the framework of evaluation study (Rita & Shokrpour, 2011). Basically, this study applied three research stages: the exploration stage to develop ~~the~~ prototype, the trials of the prototype to improve into a model and revision of the model. The research was conducted in the Faculty of Islamic Education and Teaching (FITK) of Islamic State Institute (IAIN) Surakarta from January to November 2017. As the instrument was developed under the procedures of Delphi method, we called the product as Delphi Internal Quality Assurance (henceforth is abbreviated as DIQA).

#### 3.2. Participants

This study involved 222 participants, consisting of academic experts, management of six study programs, lecturers, and students. The participants were recruited using purposive sampling techniques, and they were involved in three stages: exploration, preliminary testing, main field-testing, and ~~main~~ operational ~~main~~-testing. Table 2 describes the participants.

Table 2. Participants and their roles in the research stage

No	Research stages	Participants				Total
		Experts	Lecture	Management	Students	
1	Exploration	2	5	5	10	22
2	Preliminary testing, Delphi method	10	10	10	NA	30
3	Main field testing	-	10	10	50	70
4	Operational main testing	-	10	10	100	120
						222

#### 3.3. Research Procedures



Procedures of research indicated data collection in four stages of the research. In the exploration stages, data was collected through observation and interview with experts, lecturers, management and students in a restricted format. The aim was to identify needs, weaknesses, and expectation of how internal assurance was formatted. This stage revealed needs analysis, a prototype of the instrument for internal quality assurance and criteria of evaluation.

In the preliminary testing, the prototype of the instrument was evaluated and improved through the Delphi method in two rounds. This The Delphi method is an established research methodology aimed specifically at exploring the expected future of novel and evolutionary phenomena. The technique obtains a group of experts' most reliable consensus of opinion (Sekayi & Kennedy, 2017) by allowing them to express their own views on a topic, while taking into account the other participants' views by means of controlled feedback. The method is based on the premise that well-informed individuals, drawing on their insights and on prior experience, are better equipped to predict the future than theoretical approaches or extrapolation of trends (Grobbelaar, 2007). The responses to the questionnaires are anonymous. Participants are also provided with a summary of opinions from a previous round before answering the next questionnaire. It is believed that such a consensus process will converge the group toward the 'best' response.

In the first round, the prototype was submitted to 30 members-participants participating in the Delphi. Along with the submission, a questionnaire to assess the quality of the prototype was attached. The members also provided comments that identified the weakness of the prototype and the ways it was improved. All comments and suggestions from the experts were used as the main data to improve the prototype. In the second round, members of the Delphi method were invited to discuss the results obtained in the first round. The results were a consensus that after the improvement, the prototype was feasible to define the model of internal quality assurance for Islamic university Universities.

In the main field and operational testing, data was focused on the results of DIQA using statistical analysis using SPSS to provide evidences of validity, reliability, and conformity of each item of the DIQA, analysing the construct using Confirmatory Factor Analysis (CFA), and hypothesis testing model. Tests Test of the suitability of model, validity and reliability of this construct are statistically analysed using CFA with the help of LISREL 8.70 program (Ghozali and Fuad, 2005, pp. 29-34). The test includes include:

- 1) Goodness of Fit Index (GFI), an Index that describes the overall suitability of the model compared to the actual data. The GFI value > 0.90 suggests good suitability.
- 2) Root Mean Square Error of Approximation (RMSEA) to indicate the fit model with value < 0.05. RMSE 0.08 to 1.0 sufficient fit.
- 3) Normed Fit Index (NFI), a comparative measure between of the proposed model and the null model. The recommended value is NFI > 0.90.

Techniques of data analysis in this section are summarized in Table 3.

Table 3. Summary of Techniques of data analysis

Techniques of analysis	Application
Descriptive statistics using	Mean, percentage and determining criteria obtained from expert validation, Delphi technique, instrument sheet scores from

---

ing SPSS 17 and CFA using Lisrel program  
*ch Alpha* using SPSS 17

[a reviewer](#)  
 ct validity testing instrument of DIQA obtained from main field testing and [main](#) operational ~~main~~-testing  
 lity testing of the instrument obtained from main field-testing and [main](#) operational ~~main~~-testing.

---

#### 4. Results and Discussion

This study contributes three findings: the results of the development of the prototype of DIQA, trials of DIQA, and statistical analysis to achieve the final product of DIQA.

##### 4.1. The Development of the DIQA Prototype

This longitudinal research was conducted from January to October 2017. In general, the research procedures cover four main actions: exploration and developing the initial draft of the prototype, preliminary testing using the Delphi method, [and](#) main and operational field-testing. The instrument was developed under the Delphi method, so we named it the Delta Internal Quality Assurance (DIQA). CIPP (context, input, process, product) was used as the model of evaluation in the DIQA.

Preparation of the development of the prototype of the instrument of internal quality assurance for Islamic higher education began from the exploration process from which we developed the needs analysis. The exploration assessed the available quality instruments available and confirmed through interviews ~~with six~~ [with six](#) staff members and eight leaders at IAIN Surakarta. Documents pertaining to curriculum, quality assurance, staff, students, and implementation of overall management assurance in six study programs were evaluated.

The results approved need analysis and an initial draft to develop the prototype of DIQA. The needs analysis covers seven components that CIPP evaluation should canvas: (1) Vision and mission of the program, (2) Curriculum, (3) Competency of lecturers and administration staff, (4) Infrastructure and facilities, (5) Teaching learning process, (6) Students atmosphere supervision, and (7) Graduate learning outcomes. Preliminary, we developed ten questionnaires and 477 final items, criteria of evaluation, and scoring techniques. ~~This way, a~~ [The, the](#) two-round Delphi technique was used. ~~The, the~~ result of this stage agreed to receive the draft as a prototype of DIQA.

[Coverage](#) of the prototype appears in Table 4. Results of the preliminary test using the Delphi method are summarized in table 5 for general appearance and Table 6 to evaluate the content coverage. Criteria to evaluate the results appear in Table 7.

**Table 4. Evaluation object and number of items in DIQA**

No	Evaluation object	Number of items
1	Vision and mission document	7
2	Curriculum document	12
3	Lecturer competency	85
4	Competency of staff and administration	32
5	Infrastructure and Facilities	184
6	Document of teaching-learning planning	15
7	Teaching learning process	26
8	Assessment of teaching and learning	12
9	Document of students supervision	30
10	Graduate competency	77
Total		<b>480</b>

**Table 5. General Appearance of DIQA**

Indicator	Score	%	Criteria
Overall appearance	33	94	Very good
Clarity of evaluation	32	91	Very good
Description of the components	30	86	Very good
Practicality	32	91	Very good
Easy to understand	30	86	Very good
Very systematic	34	97	Very good
Easy to use	31	89	Very good
Clarity of writing	32	91	Very good
Selection, font, and spacing	33	94,3	Very good
Thickness of pages	30	86,7	Very good
Ability to answer	29	83	Very good
Effectiveness to do	30	86	Very good
Completion achievement	32	91	Very good

Of 24 aspects, seven reach a percentage above 76% (very good), and for aspects: page thickness, evaluation guide, time to work, implementation evaluation, data analysis, criteria determination, and evaluation report ~~reporting~~ above 51%. In terms of practicality, DIQA is categorized as practical (85% ease and 89% benefit). DIQA is categorized as efficient (86%) and deeper than AMI (89.5%). ~~Of the~~The seven internal quality standards of DIQA are divided into six books, namely: (1) Book 1: questionnaire document evaluation of mission vision, curriculum, lesson planning and student coaching document; (2) Book 2: questionnaire of lecturer competence; (3) Book 3: questionnaire of employee competency evaluation; (4) Book 4: questionnaire evaluation of facilities; (5) Book 5: questionnaire of process evaluation and assessment of learning, and (6) Book 6: questionnaire of graduate competency evaluation. Separation into six books aims to make the evaluation more efficient if it is based on the evaluation objective and reduces the "bold" impression on previous packaging evaluation models.

**Table 6. Evaluation of Instrument Dimension**

Component	Dimension	Score	%	Criteria
A. Vision and mission of study program	1. Vision and Mission of Islamic Study program	33	94,3	Very good
B. Curriculum	1. Curriculum of Islamic education	32	91	Very good
C. Competency of lecturers and Administration staff	1. Lecturer competency	27	90	Very good
	2. Competency of Administration staff	32	91	Very good
D. Infrastructure and facilities	1. General standard	34	97	Very good
	2. Mosque	30	86	Very good
	3. Classroom	29	83	Very good
	4. Library	31	89	Very good
	5. Laboratory	24	79	Good
	6. Leader room	25	80	Good
	7. Lecturer room	34	97	Very good
	8. Administration room	33	94	Very good
	9. Toilet	33	94	Very good
E. Teaching	1. Teaching-learning plan	33	94	Very good

learning Process	2. Teaching implementation	34	97	Very good
	3. Assessment	29	83	Very good
F. Students guidance	1. Aims and objectives	32	91	Very good
	2. Kinds of students guidance	29	97	Very good
G. Graduate competency	1. Personality competency	28	93	Very good
	2. Pedagogic competency	32	91,4	Very good
	3. Professional competency	29	96,7	Very good
	4. Social competency	27	90	Very good

Table 7. Degree of component in DIQA

Percentage	Criteria
1% - 20,99%	Very poor
21% - 40,99%	Poor
41% - 60,99%	Moderate
61% - 80,99%	Good
81%-100%	Very good

## 4.2. Main Field Testing and Operational Main Testing

### 4.2.1. Descriptive analysis of DIQA

The descriptive analysis is used to report [quantitatively](#) the result of rates for the DIQA [quantitatively](#). [The results](#) Result of analysis appears in Table 8.

Table 8. Descriptive analysis of instrument in Main Field Testing

Parts to be evaluated	Mean	%	Criteria
and mission of study program	3,826	1,65	Very good
structure and facilities	3,826	1,65	Very good
er competency	3,315	1,88	Very good
istration staff competency	3,681	1,03	Very good
ulum	3,826	1,65	Very good
t supervision	3,826	1,65	Very good
ient of teaching-learning process	3,826	1,65	Very good
ment	3,826	1,65	Very good
ng learning process	3,536	1,40	Very good
ite competency	3,674	1,50	Very good

Evidence shows all instruments reached >76% meaning the DIQA is "very good". Substantially, improvements are required as suggested by respondents as follows: (1) Answers to questions should be made in options a, b, c, and d, so that the respondents can easily cross-mark option they deem fix, (2) For factual data, respondents are expected to provide brief description to meet factual evaluation data; (3) Self-assessment for students should include options students can select, (4) Indicators of research and devotion of pedagogical competencies are inserted into professional competence; (5) Three indicators of learning planning are made into two only; (6) Comfort and space security indicators are added for the items of infrastructure and facilities.

Table 9. Result of quality evaluation of all study programs in FITK IAIN Surakarta

Evaluation no	Name	Mean	Category
Input	1 Vision and mission	3,83	Very good
	2 Competency of lecturer and administration staff	3,31	Very good
	3 Curriculum	3,66	Very good
	4 Infrastructure and Facilities	3,54	Good
	<b>Mean of Input</b>	<b>3,36</b>	Very good
	5 Teaching learning process	3,73	Very good
Output	6 Supervisory	3,58	Very good
	<b>Mean of Process</b>	<b>3,65</b>	Very good
	7 Competency of graduate	3,34	Good
	<b>Mean of output</b>	<b>3,34</b>	Good
	<b>Mean of Evaluation</b>	<b>3,45</b>	Very good

In summary, the average result of the quality evaluation of the study program in FITK is "excellent" with a score of 3.45. Between input, process, and output evaluation in the DIQA model, there has a comprehensive linkage, proving that good output is also determined by excellent input and process quality.

Of 24 aspects evaluated by DIQA, 20 reached above 76% (excellent), and five that include page thickness, evaluation guidance, data analysis, criteria determination, and preparation of evaluation report reached above 51% (good). Refer to table 9.

Table 10

Result of review of all instruments for PBA, PBI, PGMI, and TBI

GENERAL FORMAT				
Indicator	Max	Score	%	Criteria
Appearance & appearance of the model	54	57	105,94	Very attractive
Quality of writing	54	51	94,27	Very good
Arrangement of words, font, and spacing	54	48	88,79	Very good
Systematic	54	48	88,58	Very good
Page use	50	44	88,78	Very good
Thickness of page	54	47	87,97	Fairly thick
Readability	54	52	96,96	Easy to read
Easy to understand	54	47	87,92	Easy to understand
SUBSTANCE OF THE MODEL				
Instruction guide	50	49	98,99	Easy to understand
Depth of evaluation	50	46	92,48	Depth of evaluation already covered
Depth of component description	54	44	81,28	Component has been described
Facility to do the instrument	54	52	96,41	Easy to understand
Effort to work	54	47	87,56	Easy to do
Effort to work	50	41	82,35	Time effectiveness
Practicality	54	43	79,01	Very useful
Depth of evaluation	54	20	37,04	Used to evaluate school
Depth of evaluation	54	43	79,47	Capable to evaluate study program
Comparison to internal quality audit (AMI)	44	35	79,55	Easy to use
Comparison to other evaluation model	48	33	68,75	Easier to use
PROCEDURE OF EVALUATION				
Preparation and planning	52	18	34,62	Practical

on of evaluation	56	53	7,4	Very easy to do
s of evaluation data	52	29	2,8	Easy to do
n of evaluation criteria	52	17	0,1	Easy to do
to result of evaluation	48	15	0,3	Easy to do

In general, this means that DIQA shows that it is (92%) and efficient (86.7%) as compared with AMI. Further [testingtest](#) is required to see the validity and reliability empirically.

Commented [ST4]: Noun is missing here.

#### 4.2.2. Results of *Confirmatory Factor Analysis*

Extensive operational trials were conducted in four study programs: PBA (Arabic Study Program), PBI (English Language Education Study Program), PGMI (Teacher Education for Elementary Islamic School), and TBI (Unit of Indonesian Language Program), Faculty of Education and Teaching IAIN Surakarta involving 242 respondents.

##### 1) CFA Test to Vision and Mission Aspect

Vision and mission section consists of the vision and mission of the study program and its goals. Results of CFA ~~on mission and vision~~ with [seven7](#) items are satisfied. The measurement model achieves [a good fit](#): CFI = 0.97 and value > 0.9. ~~which means goodness of fit well.~~ In addition, t values for all items are greater than 1.00, which means that they are generally compatible with the mission aspect theory. This indicates that the [seven7](#) points are valid statement points for construct measurement from ~~aspect~~ [the aspect](#) of vision and mission of study program.

##### 2) CFA Curriculum Aspects

The curriculum aspect consists of curriculum design and curriculum criteria, measured by 12 items. Results of CFA indicate the value of CFI = 1.00 and a value > 0.9 which means DIQA model [is has](#) a good ~~of~~ fit. The t values of all items greater than 1.96 means all items in general conform to curriculum construct in the items.

##### 3) CFA aspects of Competencies of Lecturers and Administration Staff

Seven points measure [the](#) competency of lecturer and administration staff. Value of the good fit is  $p = 0.31018$  ( $p > 0.05$ ), which means these items, are good to measure. In addition, the t values for all items that are greater than 1.96, means that the items are generally compatible with the theory competency of lecturer and employee. Evidently, the items are valid for the construct measurement of lecturer and staff competencies.

##### 4) CFAforInfrastructure and Facilities

Nine points, namely general [sarpras](#), mosque, space, library, laboratory, leadership room, faculty room, administration room, and toilet, measure measurement of CFA for infrastructure and facilities. Results of test using GFI = 0.94, AGFI = 0.93, NFI = 0.94 and CFI = 1.00 are value > 0.9. This means the items [have the goodness of fit are a good fit](#). In addition, the results of t values for all items are greater than 1.96, which means all items are generally in conformity with the infrastructure stated in the DIQA model; the items contain valid statement points for the construct measurement of infrastructure aspects.

Formatted: Highlight

##### 5) CFA Model Quality Evaluation Process Aspects of Learning

The learning process has three components namely, planning, implementation and assessment and it contains 12 items. Respectively, results of test using GFI = 0.94, AGFI = 0.93, NFI = 0.94 and CFI = 1.00 are value > 0.9, meaning the items have the goodness of fit. In addition, t-values for all items greater than 1.96 means that the items are generally compatible with the ethical and regulatory aspects of [the](#) theory. The items are valid statement points for construct measurement of aspects of the learning process.

#### 6) CFA Aspects of Student Development

Aspects of student coaching consists of guidance, guardianship, skill practice, literary reading and bilingual are measured with 30 items of [the](#) statement. CFA results indicating GFI = 0.97, AGFI = 0.97, NFI = 0.93 and CFI = 0.97 have a value > 0.9. This means the items in DIQA have a goodness of fit. The t-values for all grains that are valued at greater than 1.96, means the grains are in conformity with the theory of student coaching aspects. The items are valid statement points for the construct measurement of student coaching aspects.

#### 7) CFA Aspects of Graduate Competence

The CFA measurement for the graduate competency consists of 4 components, namely personality competency, pedagogical competence, professional competence and social competency component and is measured by 15 items. The results CFA are GFI = 0.97, AGFI = 0.97, NFI = 0.93 and CFI = 0.97 with a value > 0.9. It means the items ~~have the goodness of fit~~ are a good fit. In addition, the t values for all grains that are greater than 1.96, ~~which~~; means the grains are generally in conformity with the theory of student coaching aspects. The items are valid statement points for the construct measurement of student coaching aspects.

#### 4.2.3. Result of Hypothetical Model Testing DIQA

The modelling hypothesis testing of DIQA shows evidence that DIQA necessitates all components of input, process, and output evaluation to be evaluated. Evaluation of input quality will determine the quality of [the](#) process, and quality of the process affects the output.

**Input Quality Evaluation.** This evaluation is to examine the quality of Islamic education study programs by looking at: (a) vision and mission of the study program; (b) the curriculum and curriculum design; (c) competence of lecturers and employees related to pedagogical competence, professional competence, social competence and personality competence; and (c) infrastructure and facilities, such as mosques, classrooms, libraries, multimedia, laboratories, leadership rooms, faculty rooms, administration rooms and toilets.

**Quality Process Evaluation.** This evaluation is to assess the quality of the study program by looking at: (a) implementation of learning related to the planning, the implementation process of learning and assessment; (b) student coaching related to thesis guidance, guardian of study, Al-Qur'an literacy coaching, expertise practice, and language development of students.

**Evaluate Output Quality.** This evaluation is to see the quality of graduates by measuring teacher professionalism that includes pedagogical competence, professional competence, social competence and personality competence.

The statistical analysis results are clarified below. The DIQA test with CFA using SEM, proves that DIQA has good ability in matching data (good fit). Evidences on the standardized loading of a hypothetical model of component relations, variables of input quality evaluation, process quality and output quality, show that correlational indicators among variables have a high loading factor  $\geq 0.3$  (Tabachnick&Fidell, 2007: 217; Donna, 2009: 215). This means that the main indicator of latent construct of DIQA model has been well rated and understood by the respondents. The DIQA model constructs are well applied and highly deserve to be maintained and used.

The loading factor value means (1) quality evaluation of input to process quality has a loading factor value of 0.32 with the quadratic value of  $0.32^2 = 10.24$ . This means that a 10.24% variant of input quality influences process quality, and (2) evaluation of process quality to output quality has a loading value equal to 0.57 with the quadratic value ( $0.57^2 = 0,3249$ ). Also, 32,49% variant of process quality evaluation influences output quality. Thus, evaluation of the quality of input affects the quality of the process and ultimately contributes quality evaluation of teacher quality output. This result is reinforced with t-value with cut-off 5% (value  $t = 1.96$ ). The evidence implies (1) quality of input significantly influences process quality, (2) evaluation of process quality significantly influences output quality.

## 5. Conclusion and Implication

In summary, this study contributes three findings: kind of needs for the internal quality assurance, the process of development of DIQA and empirical evidence to achieve the final product of a DIQA model.

- 1) Needs analysis is useful for internal quality assurance in the Islamic university including: vision and mission of study program, curriculum, competency for lecturer and administration staff, infrastructure and facilities, students coaching, teaching-learning process, assessment for teaching-learning process, and graduate competency. DIQA has ten different sets of questionnaires and total items 480. DIQA is based on the CIPP to develop the constructs, method of evaluation, and procedures to include research procedures.
- 2) The development of items is begun from the exploration stage to prepare the needs analysis and initial package of the DIQA. Validation of items starts from the preliminary testing through the Delphi method conducted in two rounds. The results of the Delphi method are used to improve the prototype into a model. Revision addresses that DIQA has seven dimensions of evaluation, ten questionnaires, and 477 items as a strong format. Inspired by the Delphi method, the model is named DIQA that stands for Delta Internal Quality Assurance.
- 3) DIQA has served to improve for qualitative and quantitative techniques. A ~~Qualitative~~ qualitative approach has been used to develop the prototype during preliminary testing and main field-testing to improve the model of DIQA. Quantitatively, statistical analysis using SPSS and ISREL is prepared to see evidence of validity, reliability, items, and options in each questionnaire. The results evidently prove that the appearance of DIQA is very good, individual items are valid, and consistency to evaluate the study program is reliable. The final version of DIQA approves to consist of seven dimensions of internal quality assurance, ten kinds questionnaire, and 477 items, improving questionnaire from AMI and BAN-PT.



This research, however, experienced limitations: less cooperative respondents, uncertain timing of the evaluation in the study program, a dilemma to determine respondents for measurement of the competence graduates, and positioning DIQA with government accreditation. Less cooperative respondents have made the objectives of the evaluation are not clear and not fully match. Time of evaluation frequently does not conform so that the external validity may shed. Determining student graduates who actually represent graduate competence is frequently more subjective. Tracer studies that serve comprehensive data on graduate competency are not fully available. Finally, DIQA cannot relocate government evaluation for accreditation.

The limitations give implications for the study program and future research. First, benchmarking is the ultimate goal of accreditation. The benchmarking through accreditation makes the study program strive to get a good accreditation value meeting all indicators required in accreditation. Thus, the Islamic education program must follow the rules or standards set by BAN-PT, though a specificity and peculiarity of Islamic values are not covered.

It implies that the study programs can make use of DIQA to accommodate Islamic values, as the internal quality assurance instrument. DIQA serves the internal quality evaluation for the Islamic education program in producing professional teachers and helps prepare the external evaluations, such as accreditation. DIQA is evidently an effective way to know the quality of Islamic education programs in producing professional teachers because it is able to: 1) produce a study profile with up to date data and information; 2) planning and improving the quality of Islamic education programs on a regular basis; 3) provide information about the community and parties in need; and 4) prepare the study program in the face of external evaluation or accreditation.

The second implication that arises pertains to the future research. This study has obtained evidence that the reluctance of respondents to be involved in the research and uncertain time to conduct the evaluation affect the validity of the items. This implies that DIQA contains less comprehensive attributes of accreditation and aligning DIQA with BAN-PT is something problematic. Future research can verify the items of DIQA are improving the items and dimension of evaluation. Efforts to align DIQA as initial accreditation training are recommended for use by the Islamic universities.

#### **REFERENCES** ---will be revised based on the clean version after final edit.

Abadie-Mendia, Trudy, "Accountability and Quality in Higher Education: A Case Study" (2013). *UNF Theses and Dissertations*. 375. University of North Florida.  
<http://digitalcommons.unf.edu/etd/375>

Arcaro, S.J. (1995). *Quality in Education: An Implementation Handbook*. Alihbahasa Yosallriantara. Yogyakarta: PustakaPelajar

BAN-PT (National Accreditation Body for Higher Education). (2002). *Guidelines for Internal Quality Assessment of Higher Education*. Second edition. Ministry of National Education.

Bridge, R.G. Judd, C.M. & Moock, P.R. (1979). *The Determinants of Educational Outcomes*, Massachusetts: Ballinger Publishing Company

- Chinta, Ravi; Kebritchi, Mansureh&Ellias, Janelle. (2016). A conceptual framework for evaluating higher education institutions.*International Journal of Educational Management*, Vol. 30 No. 6, pp. 989-1002. DOI: 10.1108/IJEM-09-2015-0120.
- Choiriyah, Siti. (2018). *Model Evaluasi Mutu Internal Program Studi Pendidikan Islam dalam Menghasilkan Guru Profesional*. Disertasi Doktor Tidak Diterbitkan. Yogyakarta: Program Studi Penelitian dan Evaluasi Pendidikan, Universitas Negeri Yogyakarta.
- Directorate General of Higher Education. (2003). Higher Education Long Term Strategies: *HELTS 2003 - 2010*. Department of National Education, Jakarta.
- Donna. H. (2009). *Confirmatory factor analysis*. Oxford: Oxford University Press.
- Ewell, Peter. 2010. "Twenty Years of Quality Assurance in Higher Education: What's Happened and What's Different?" *Quality in Higher Education* 16 (2): 173–75.
- Fitri, Zainul Agus. (2016). Quality Assurance System between the Islamic University and the State University. *Aljamiah*. Vol. 2 No. 2, pp. 208-235. DOI: <http://dx.doi.org/10.15575/jpi.v2i2.787>.
- GAO (US Government Accountability Office). (2014) and Andrea Fuller and Douglas Belkin, "The Watchdogs of College Education Rarely Bite," *Wall Street Journal*, June 17, 2015, <https://www.wsj.com/articles/the-watchdogs-of-college-education-rarely-bite-1434594602>.
- Gaspersz, V. (2005). *Total Quality Management*. Jakarta: PT Gramedia Pustaka Utama
- Ghozali, I. (2009). *Aplikasi Analisis Multivariate dengan Proses SPSS*. Semarang: Badan Penerbit Universitas Diponegoro
- Grobelaar, S. (2007). *Data Gathering: Delphi Method*. R&D in the National System of Innovation: A System, Dynamic Model. University of Pretoria Etd. Pretoria, South Africa.
- Haris, Ikhfan. (2013). Assessment on the Implementation of Internal Quality Assurance at Higher Education (An Indonesian Report). *Journal of Educational and Instructional Studies in the World*. Vol. 3 No 4, pp. 45-49.
- Harvey, L. and Green, D. (1993) Defining quality. *Assessment and Evaluation in Higher Education*, 18 (1): 9-34.
- Kelchen, Robert. (2017). *Higher Education Accreditation and Federal Government*. Washington DC, Urban Institute. Urban.org.
- Koslowski III, F.A. (2006). Quality and assessment in context. *Quality Assurance in Education*, 3, 277-288. doi: 10.1108/09684880610678586

- Lenn, M. P. (2004). *Quality assurance and accreditation in higher education in East Asia and the Pacific* Washington, DC: The World Bank.
- MSCHE (Middle States Commission on Higher Education). 2014. "Standards for Accreditation and Requirements of Affiliation." 13th ed. Philadelphia: MSCHE.
- NEA (National Education Association). (2006). *School quality: Issues in education*. (<http://www.nea.org/jobs/>). Diambil dari <http://www.proquest.com/pqdweb.html>
- Patil, S. Arun & Pudlowski, J. Zenon. (2005). Important Issues of the Accreditation and Quality Assurance and a Strategy in the Development of Accreditation Framework for Engineering Course. *Global Journal of Engineering Education*. Vol. 9 No. 1, pp. 49-62.
- Podsakoff, P.M., MacKenzie, S.B., Paine, J.B. and Bachrach, D.G. (2000), "Organizational citizenship behaviors: a critical review of the theoretical and empirical literature and suggestions for future research." *Journal of Management*, Vol. 26 No. 3, pp. 513-563.
- Rich, B.L., Lepine, J.A. and Crawford, E.R. (2010), "Job engagement: antecedents and effects on job performance." *Academy of Management Journal*, Vol. 53 No. 3, pp. 617-635.
- Rita, Rezaee & Shokrpour, Nasrin. (2011). Performance Assessment of Academic Departments: CIPP Model. *European Journal of Social Sciences*, Vol. 23 No. 2. pp. 228-236.
- Roffe, I. M. (1998). Conceptual problems of continuous quality improvement and innovation in higher education. *Quality Assurance in Education*, 6, 74-82.
- Rossi, P. H., Lipsey, M. W., & Freeman, H. E. (2004). *Evaluation: A systematic approach* (7th ed.). CA: Sage Publications
- Sallis, E. (2002). *Total quality management in education*. London: Kogan Page Limited
- Sekayi, D., & Kennedy, A. (2017). Qualitative Delphi Method: A Four Round Process with a Worked Example. *The Qualitative Report*, 22(10), 2755-2763. Retrieved from <http://nsuworks.nova.edu/tqr/vol22/iss10/15>
- Shafer, B. S., & Coate, L. E. (1992). Benchmarking in higher education: A tool for improving quality and reducing cost. *Business Officer*, 26(5), 28-35.
- Shakhnoza (2009). *Perception Of Quality In Changing University Education In Kazakhstan*. Yakubova: Kent State University College and Graduate School of Education, Health, and Human Services

- Tabachnick, B. G., & Fidell, L. S. (2007). *Experimental designs using ANOVA*. Colombia: Thomson Brooks/Cole
- Tam, Siu-ling, Maureen. (2001). *Assessing Quality in Higher Education by Examining the Effects of University Experiences on Learning Outcomes and Student Development*. Ph.D. Dissertation. University of Durham. Stockton, Durham City, UK.
- Utuka, Godwin. (2012). *Quality Assurance in Higher Education: Comparative Analysis of Provision and Practices in Ghana and New Zealand*. Ph.D. Dissertation. Wellington: Victoria University of Wellington.
- Van Vught F.A., & Westerheijden D.F. (1994). Towards a general model of quality assessment in higher education. *Higher Education*, 28, 355-371.
- Watty, K. (2006). Want to know about quality in higher education? Ask an academic. *Quality in Higher Education*, 12(3), 291-301.
- Woodhouse, D. (1999). Quality and quality assurance: an overview. In H. De Wit & J. Knight (Eds.), *Quality and internationalisation in higher education* (pp. 29-43). Paris: OECD.