Developing Delta Internal Quality Assurance to Evaluate
the Quality of Indonesian Islamic Universities

Siti Choiriyah[1], Kumaidi[2], Badrun Kartowagiran[3]

Abstract

The purpose of this study is to examine aspects of internal quality assurance to evaluate Indonesian Islamic universities, develop Delta Internal Quality Assurance (DIQA), and provide empirical evidence for using DIQA as a standard model of evaluation. It is a research and development (R&D) endeavor in the context of the input, process, and product. This study applies four cycles, namely exploration, preliminary testing, main field testing, and main operational testing, and the development process took place through a two-round Delphi method. In terms of findings, the study presents a prototype instrument named DIQA, which is subsequently refined though the main field and operational testing phases. It is then processed using CFA and LISREL statistical analysis to determine its validity and reliability. The final version of DIQA accommodates seven dimensions of evaluation, 10 sorts of questionnaire, and 477 question items. The implication is that DIQA will help face the challenge of aligning with the national accreditation system set by the government. It follows that DIQA should be publicly and disseminated, and Islamic universities should be encouraged to use DIQA confidently. In terms of originality and value, DIQA has the special property of promoting Islamic values in the internal quality assurance, which is important to the specific properties of Islamic universities in Indonesia.

Keywords–internal quality, benchmarking, quality assurance, Islamic values.

Introduction

This article derives from a doctoral dissertation reporting on the development of an instrument for internal quality assurance for Islamic universities in Indonesia that has a focus on vision and mission, curriculum, teaching/learning process, infrastructure and facilities, and student outcomes. The instrument was developed using the Delphi method, and we named it Delta Internal Quality Assurance (or DIQA for short). This study was conducted for three reasons: i) because quality assurance is central to higher education management (Legčević and Hećimović, 2016; Tam, 2001) and part of benchmarking (Shafer and Coate, 1992); and ii)

[1]Dr Candidate. State Islamic Institute of Surakarta, Indonesia. Email: choiriain@gmail.com
[2]Prof. Dr. Muhammadiyah University of Surakarta, Indonesia. Email: kum231@ums.ac.id
[3]Prof. Dr. State University of Yogyakarta, Indonesia, Email: badrunkw@yahoo.com

DIQA is an attractive tool for evaluating internal quality assurance in Islamic universities (Choiriyah, 2018).

The issue of the quality of higher education and the need to apply continuous quality improvement (CQI) are obvious characteristics of a good education development policy (Legčević and Hećimović, 2016). The key to quality assurance is to inform the academic community—from teachers to students and the administrative staff—that quality assurance will be put into practice and implemented in their respective institution (Tam, 2001).

Higher education institutions have recently put emphasis on academic quality as a benchmarking process, introducing systematic evaluations of education at departmental, faculty, and university-wide levels (Rossi, Lipsey and Freemans, 2004). Higher education, much like any industry, has to benchmark its operations with set performance targets. This benchmarking process helps to understand the drivers of the processes and their output quality (Shafer and Coate, 1992). Benchmarking provides institution managers with an external reference or standard for evaluating the quality and value of the internal activities, practices, and processes (Tam, 2001).

The quality of higher education should be measured according to its purposes and main goals. The assessment program determines the measured outcomes and the quality-measurement approach of Tam (2001), thus emphasizing the quality of research, public service, and student education (Patil and Pudlowski, 2005; Tam, 2001).

Unfortunately, the Indonesian education system face a plethora of quality problems, specifically low input quality and process-to-output quality, school graduates' outcomes, and an in ability to meet industrial needs (Fitri, 2016). Many higher education institutions are unaccredited, and there is a dire shortage of human capital. Meeting global standards is hard, because greater investment and internationalization in research capacity is needed in universities (Woodhouse, 1999, p. 20). Islamic universities should therefore devise their own instruments to aid stakeholders in improving the culture of educational quality. This study proposes DIQA, an instrument to evaluate internal quality assurance for Islamic universities. Three research questions guide the investigation:

1. What aspects of internal quality assurance are needed to evaluate Indonesian Islamic universities and help them to operate standard services for university teaching programs?

2. What are the development processes of DIQA for achieving a standard evaluation model?

3. What evidence is there to support DIQA's role in assessing the internal quality of Indonesian Islamic universities?

Literature Review

Internal quality assurance

Before conceiving internal quality assurance, quality must first be defined. Watty (2006, p. 293) argues that quality is about efficiency, high standards, excellence, value for money, fitness for purpose, and customer focus. Fitness for purpose includes the mission, goals, objectives, and specifications, and it means an organization uses procedures that are appropriate and effective for its stated purposes. Table 1 shows the five components of as defined by Harvey and Green (1993).

With this in mind, quality assurance is a mechanism to control quality. It should therefore ensure that adequate conditions and provisions are in place to enable students to reach a certain standard. Woodhouse (1999, p.30), meanwhile, defines quality assurance as follows: "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."

Table 1.

Harvey and Green's 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 good value for their investment
Quality as transformation	A focus on ensuring that students are genuinely empowered as a result of their learning

According to Lenn (2004), approaches in quality assurance vary in accreditation, assessment, academic audit, and external examination, each of which allows for the development and application of criteria for a program's set standards by the accrediting body. The motivation

may be to identify further improvements for a program or the larger educational system. In addition, Arcaro (1995, p. 1) suggests that a quality program basically includes a commitment to change, a good understanding of the condition of the program or institution, a clear vision of the future that everyone sticks to, and plans for implementing educational quality.

Practices inquality assurance

Practices in quality assurance relate to assessment and benchmarking. Pressure to facilitate universal access makes the assessment of higher education a major public concern (Tam, 2001; Patil and Pudlowski, 2005). Koslowski (2006) suggests that much like with industry, higher education is a measurable product, service, or knowledge that is essentially evaluated by customer satisfaction. A university's quality is therefore determined by its output, such as whether or not it uses its resources efficiently to produce highly skilled, employable graduates. Quality is defined by the customer, and management is responsible for achieving that quality.

Koslowski (2006) asserts that a process has quality when higher education institutions view it as valuable, measurable, and improvable. Assessment is a measurable process to help improve quality by evaluating it. Assessments can take the form of guided self-assessment, intermediary conduct assessment, independent self-assessment, and student-competencies-based assessment. Guided self-assessment can be peer-review based, similar to business certification standards like ISO 9000. Koslowski (2006) believes that the academic audit has become the dominant model for institutional assessment in higher education. Through independent self-assessment, institutions assess the needs of its customers, resulting in a process of education.

In addition, internal quality assurance (Utuka, 2012) includes (i) a quality assurance policy that is publicly available and part of a strategic planning (Tam, 2001); (ii) the design and approval of programs to meet the set objectives and achieve the intended learning outcomes, such as through regulation of student admission, progression, recognition, and certification (Tam, 2001; Patil and Pudlowski, 2005); (iii) teaching staff, learning resources, and student support; (iv) effective management of programs, where published information is clear, accurate, objective, up-to date, and readily accessible (Tam, 2001); and (v) periodic monitoring and

reviewing of programs to continue meeting the set objectives by responding to the needs of students and society (Utuka, 2012; Tam, 2001; Patil and Pudlowski, 2005).

To implement their internal quality assurance, universities apply their available resources and quality assessment using a continuous quality improvement (CQI) and people-oriented approach (Yakubova, 2009; Roffe, 1998). Roffe (1998, p. 74) states that CQI derives from the Japanese term *kaizen*, which means to make slow, never-ending improvements in all aspects of life. Outside the Japanese context, *kaizen* or CQI is to continuously make small improvements to an existing system through the people who manage or work in that system. The structural steps to this are as follows:

1. Define the area of improvement

2. Analyze and select appropriate problems

3. Identify causes

4. Plan countermeasures

5. Implement them

6. Confirm the results

7. Standardize (Roffe, 1998).

BAN-PT (2002) recommends that internal quality assurance be controlled through the PDCA model (plan, do, check, action) that results in CQI. PDCA-based quality control management works on the following principles.

1. Quality first: The thoughts and actions of education managers should be focused on quality.

2. Everything for stakeholders: All thoughts and actions of education managers must aim to satisfy stakeholders.

3. Our stakeholders: Any person working in higher education should aim to satisfy the consumers of their work product.

4. Speak with data: Any action and decision taken in higher education should be based on analyzing collected data rather than supposition.

5. Upstream management: All decision-making in higher education must be participatory not authoritative.

Quality assurance should be internally driven, institutionalized within each organization's standard procedure, and involve external parties that produce outputs and outcomes as part of

public account ability (BAN-PT, 2002), thus having internal and external quality assurance systems. The internal quality assurance (Haris, 2013) comprises three major areas:

1. Assessing the extent to which internal quality assurance is implemented

2. Presenting a quality profile for every learning unit to reveal the strengths and weaknesses of the quality assurance program

3. Recording feedback, suggestions, and recommendations to the university thatcan be implemented in the internal quality assurance to improve, develop, and strengthen its implementation.

Quality assurance and measurement

Quality assurance and quality measurement are used when the structure of the higher education sector become more complex (Tam, 2001; Kelchen, 2017). The internationalization process in higher education, as well as the growth in free trade, has caused universities to focus on quality and emphasize standards for assessing the quality of their education structures (Anderson, Johnson and Milligan, 2000, p. 52). Any industrial activity covering an input, process, and output can be adapted for quality assessment in three stages according to Anderson Johnson and Milligan (2000), as shown in Figure 1.
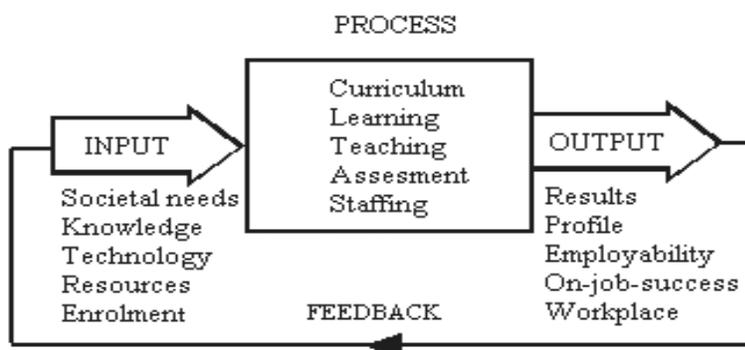

Fig. 1. *The block diagram of an educational cycle (adopted from Anderson, Johnson and Milligan, 2000, p. 52)*

Educational Input: The *input* parameters relate to student intake/enrolment into an educational process and comprise:

- Societal needs
- New knowledge

- Advancing technologies
- Human and material resources
- Student enrolment process
- Student fee structure
- Student eligibility criteria

Educational Process: The educational process lies between the input and the output, where learning is facilitated. It may comprise the following important factors:

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

Learning Outcomes: The *output* component is associated with student outcomes following the course curricula. It comprises:

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

Measurement and benchmarking in quality assurance

Measurement and benchmarking are inseparable elements of quality assurance. According to Chinta et al., "What you measure is what you get" (2016, p. 989) and "What benchmarks you use is what meaning you get" (2016, p. 990). Measurement is the basis for applying multiple metrics in performance management across multiple dimensions (Chinta, et. al. 2016; Podsakoff, et. al. 2000). Benchmarking is then evaluating an action with a standard for

comparison. Employee engagement can then be enhanced with greater shared understanding of the metrics used (Chinta, et. al. 2016; Rich, et. al, 2010; Fuller and Belkin, 2015).

With benchmarking, factors affecting the quality of education systems include effective learning and teaching, leadership, lecturers, students, institutional management, the physical environment and resources, stakeholder satisfaction, institutional culture, learning outcomes/performance, and accountability (Bridge, Judd and Moock, 1973; Abadie-Mendia, 2013). Sallis (2002) and Ewell (2010) indicate ten indicators for quality assurance: (1) effective learning and teaching, (2) leadership, (3) staff, (4) students, (5) standards, (6) organization, (7) physical environment and resources, (8) external relations, (9) access, and (10) service to customers.

Methods

Approach

This is a research and development (R&D) study that applies a qualitative and quantitative approach. It adapts the R&D approach of Borg and Gall (2003), whose steps are: (1) researching and collecting information; (2) planning; (3) developing the preliminary form of the product; (4) preliminary field testing; (5) revising main product; (6) main field testing; (7) revising operational product; (8) operational field testing; (9) revising final product; and (10) disseminating and implementing. In the development process, we use a two-round Delphi method and CIPP (context, input, process, product) as the framework of evaluation (Rezaee and Shokrpour, 2011; NEA (National Education Association), 2006). This study essentially applies the following stages: exploration to develop prototype, trials to refine the prototype into a model, and revision of that model. Research was conducted at the Faculty of Islamic Education and Teaching (FITK) of the Islamic State Institute (IAIN) of Surakarta from January to November 2017.

Participants

This study involved 222 participants (see Table 2), including academic experts, managers of six study programs, lecturers, and students. Participants were recruited using purposive sampling, and all were all involved in the exploration, preliminary testing, main field testing, and operational main testing (Borg & Gall, 2003).

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

Research procedures

During the exploration stage, data were collected through observation and interviews with experts, lecturers, management, and students in a restricted format. This stage revealed the needs analysis and a prototype for internal quality assurance and evaluation criteria.

In the preliminary testing, the prototype was evaluated and improved through a two-round Delphi method, which is an established research methodology for exploring the expected future of novel and evolutionary phenomena. It involves obtaining a group of experts' most reliable consensus of opinion (Sekayi and Kennedy, 2017) by allowing them to express their views on a topic. The method is based on the premise that well-informed individuals can better predict the future than a simple extrapolation of trends (Grobbelarr, 2007). 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 guides the group toward the "best" response.

In the first round, the prototype was submitted to 30 members participating in the Delphi method, and a questionnaire to assess the quality of the prototype was attached. The members provided comments identifying weaknesses in the prototype and suggesting ways to improve it. In the second round, members were invited to meet and discuss the results from the first round. The consensus was that the prototype was acceptable as an internal quality assurance model for Islamic universities ((Sekayi and Kennedy, 2017).

In the main field testing and operational main testing, the data were subjected to statistical analysis. SPSS was applied to provide evidence for the validity, reliability, and conformity of each item of the DIQA, using Confirmatory Factor Analysis (CFA) and a

hypothesis testing model using the LISREL 8.70 program (Ghozali, 2009, pp.29-34), as summarized in Table 3. The test includes:

1. **Goodness of Fit Index (GFI)** that describes the overall suitability of the model compared to the actual data, where a GFI value greater than 0.90 suggests good suitability.

2. **Root Mean Square Error of Approximation (RMSEA)** to indicate the fit model with value <0.05, where RMSE 0.08 to 1.0 is a sufficient fit.

3. **Normed Fit Index (NFI)** as a comparative measure between the proposed model and the null model, with the recommended value of NFI being greater than 0.90.

Table 3.

Summary of data analysistechniques

Analysis technique	Application
Descriptive statistics	Calculating mean and percentage and determining criteria obtained from expert validation, Delphi technique, and instrument sheet scores from *reviewer*
EFA using SPSS 17 and CFA using LISREL program	Construct validity testing instrument for DIQA obtained from main field testing and operational main testing
Cronbach Alpha using SPSS 17	Reliability testing of the instrument obtained from main field testing and operational main testing.

Results and Discussion

The development of the DIQA prototype

The development of DIQA began with exploring the needs analysis. The needs analysis portrays seven dimensions that CIPP evaluation requires (Gaspersz, 2005; MSCHE (Middle States Commission on Higher Education), 2005), namely (1) vision and mission of the program, (2) curriculum, (3) competency of lecturers and administrative staff, (4) infrastructure and facilities, (5) teaching and learning processes, (6) student supervision atmosphere, and (7) graduate learning outcomes. A first, we developed 10 sorts of questionnaire with 483 items, which after revision became 477 items. A two-round Delphi technique was applied to the DIQA prototype (as shown in Table 4) at this stage. The results of the preliminary format are summarized in Tables 5 and 6, and the criteria to evaluate the results are 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 student supervision	30
10	Graduate competency	77
Total		**480**

Table 5.

The general appearance of DIQA

No.	Indicator	Score	%	Criteria
1	Cover	33	94	Very good
2	Content appearance	32	91	Very good
3	Scope of evaluation	30	86	Very good
4	Depth description of the components	30	86	Very good
5	Readability	32	91	Very good
6	Ease of understanding	30	86	Very good
7	Systematic writing	34	97	Very good
8	Language use	31	89	Very good
9	*Layout* of writing	32	91	Very good
10	Word choice, font, and spacing	33	94.3	Very good
11	Thickness of pages	30	85.7	Very good
12	Practicality of answer	29	83	Very good
13	Time effectiveness	30	86	Very good
14	Evaluation achievement	32	91	Very good

Of the 24 aspects, 7 were above 76% (very good). In particular, DIQA is practical (85% ease and 89% benefit) and efficient (86%), and it goes deeper than AMI (89.5%). DIQA is divided into six books considering efficiently used, evaluation-based objectives. These books are:

- Book 1: questionnaire evaluating mission vision, curriculum, lesson planning, and student coaching
- Book 2: questionnaire for lecturer competence
- Book 3: questionnaire for employee competency evaluation

- Book 4: questionnaire for evaluation of facilities
- Book 5: questionnaire for process evaluation and assessment of learning
- Book 6: questionnaire for graduate competency evaluation.

Table 6.

Evaluation of instrument dimensions

Component	Dimension	Score	%	Criteria
Vision and mission of study program	Vision and mission of Islamic study program	33	94.3	Very good
Curriculum	Curriculum of Islamic education	32	91	Very good
Competency of lecturers and administrative staff	Lecturer competency	27	90	Very good
	Competency of administrative staff	32	91	Very good
Infrastructure and facilities	General standard	34	97	Very good
	Mosque	30	86	Very good
	Classroom	29	83	Very good
	Library	31	89	Very good
	Laboratory	24	79	Good
	Leader room	25	80	Good
	Lecturer room	34	97	Very good
	Administration room	33	94	Very good
	Toilet	33	94	Very good
Teaching learning Process	Teaching/learning plan	33	94	Very good
	Teaching implementation	34	97	Very good
	Assessment	29	83	Very good
Student guidance	Aims and objectives	32	91	Very good
	Kinds of student guidance	29	97	Very good
Graduate competency	Personality competency	28	93	Very good
	Pedagogic competency	32	91,4	Very good
	Professional competency	29	96.7	Very good
	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

Main field testing and operational main testing

Descriptive analysis of DIQA

Descriptive analysis was used to quantitatively report the result rates for DIQA, as shown in Table 8. This shows all instruments attaining greater than 76%, meaning that DIQA is "very good."

Table 8.

Descriptive analysis of the instrument in themain field testing

No.	Parts to be evaluated	Mean	%	Criteria
1	Vision and mission of study program	3.826	95.65	Very good
2	Infrastructure and facilities	3.826	95.65	Very good
3	Lecturer competency	3.315	82.88	Very good
4	Administrative staff competency	3.681	92.03	Very good
5	Curriculum	3.826	95.65	Very good
6	Student supervision	3.826	95.65	Very good
7	Document of teaching learning process	3.826	95.65	Very good
8	Assessment	3.826	95.65	Very good
9	Teaching learning process	3.536	88.40	Very good
10	Graduate competency	3.674	85.50	Very good

Table 9.

Result of quality evaluation forall study programs of FITK IAIN Surakarta

Evaluation	No	Name	Mean	Category
	1	Vision and mission	3.83	Very good
Input	2	Competency of lecturer and administration staff	3.31	Very good
	3	Curriculum	3.66	Very good
	4	Infrastructure and Facilities	3.54	Good
	Mean of Input		*3.36*	*Very good*
Process	5	Teaching learning process	3.73	Very good
	6	Supervisory	3.58	Very good

Evaluation	No	Name	Mean	Category
		Mean of process	*3.65*	*Very good*
Output	7	Competency of graduates	3.34	Good
		Mean of output	*3.34*	*Good*
		Mean of evaluation	*3.45*	*Very good*

The data in Tables 7 to 9 summarize the results of the quality evaluation, with the study programs in FITK being "very good" on average with a score of 3.45. Between the input, process, and output evaluation in the DIQA model, there is comprehensive linkage, showing that good output is also determined by excellent input and process quality.

As Table 9 suggests, of the 24 aspects evaluated for DIQA, 20 scored over 76% (excellent), with five—page thickness, evaluation guidance, data analysis, criteria determination, and preparation of evaluation report—scoring above 51% (good). In addition, table 10 shows that in general, DIQA is easy to use (92%) and efficient (86.7%) when compared with AMI. The in-depth analysis using CFA proves the result of the item index, indicating validity and reliability for each item of the DIQA questionnaire.

Table 10.

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

No	Indicator	Max	Score	%	Criteria
	General format				
1	Package & appearance of the model	464	357	76.94	Very attractive
2	Layout of writing	464	391	84.27	Very good
3	Selection of words, font, and spacing	464	408	87.9	Very good
4	Systematic writing	464	398	85.8	Very good
5	Language use	460	404	87.8	Very good
6	Thickness of page	464	277	59.7	Fairly thick
7	Readability	464	462	99.6	Easy to read
8	Easy to understand	464	427	92	Easy to understand
	Substance of the model				
1	Evaluation guide	460	349	75.9	Easy to understand
2	Scope of evaluation	460	436	94.8	Scope of evaluation already covered
3	Depth of component description	464	384	82.8	Component been described
4	Guidance to use the instrument	464	452	97.41	Easy to understand
5	Ease of work	464	397	85.6	Easy to do
6	Time to work	460	391	85	Time effective
7	Significance	464	413	89.01	Very useful
8	Urgency of evaluation	464	420	90.5	Needed to evaluate school

9	Achievement of evaluation	464	393	84.7	Enables evaluation of study program
10	Comparison to internal quality audit (AMI)	444	385	86.7	Easier to use
11	Comparison to other evaluation model	448	383	85.5	Easier to use
	Procedure of evaluation				
1	Preparation and planning	452	418	92.5	Practical
2	Execution of evaluation	456	353	77.4	Very easy to do
3	Analysis of evaluation data	452	329	72.8	Easy to do
4	Decision of evaluation criteria	452	317	70.1	Easy to do
5	Report result of evaluation	448	315	70.3	Easy to do

Results of confirmatory factor analysis

1) CFA test for vision and mission aspect

The vision and mission section, as the name suggests, comprises the vision and mission of the study program and its goals. Results of the CFA on mission and vision shows seven items are satisfied. The measurement model achieves a good fit, as demonstrated by a CFI of 0.97, which is greater than 0.9. In addition, the t-values for all items are greater than 1.00, which means they are generally compatible with the mission aspect theory. This indicates that the seven points are valid points for constructing a measurement of the vision and mission of a study program.

2) CFA curriculum aspects

The curriculum aspect comprises curriculum design and curriculum criteria, as measured by 12 items. The CFA results indicate a CFI value of 1.00. Being greater than 0.9, this means the DIQA model achieves a good fit. The t values of all items are greater than 1.96, meaning all items generally conform with the curriculum construct.

3) CFA for aspects of competency for lecturers and administrative staff

The competency of lecturers and administrative staff is measured according to seven points. The value of good fit is p = 0.31018 (p> 0.05),meaning these items are suitable. In addition, the t values for all items are greater than 1.96, so the items are generally compatible with staff competencies and valid for measuring them.

4) CFA for infrastructure and facilities

Measurement of CFA for infrastructure and facilities is performed according to nine points: general standards, mosque, space, library, laboratory, leadership room, faculty room, administration room, and toilet. The test yielded the following results: GFI = 0.94, AGFI = 0.93, NFI = 0.94 and CFI = 1.00 (> 0.9). This means the items have a good fit. In addition, the t values

for all items are greater than 1.96, which means all items generally conform to the infrastructure stated in the DIQA model, and they contain valid points for measuring infrastructure aspects.

5) CFA for aspects of learning

The learning process covers three components—namely planning, implementation and assessment—and comprises 12 items. The test yielded the following results: GFI = 0.94, AGFI = 0.93, NFI = 0.94 and CFI = 1.00 (> 0.9). The items therefore have a good fit. In addition, the t-values for all items are greater than 1.96, meaning that they are generally compatible with aspects of learning and form valid points for measuring the process.

6) CFA for aspects of student development

Aspects of student coaching cover guidance, guardianship, skill practice, literary reading, and bilingual support, as measured by 30 items. The test gives the following result: GFI = 0.97, AGFI = 0.97, NFI = 0.93 and CFI = 0.97 (> 0.9). This means DIQA's items are a good fit here. The t-values for all items are also all greater than 1.96, meaning they conform with student coaching aspects and are valid points for measuring them.

7) CFA for aspects of graduate competence

Graduate competency comprises four components, namely personality competency, pedagogical competency, professional competency, and social competency components, as measured by 15 items. The results of the CFA were: GFI = 0.97, AGFI = 0.97, NFI = 0.93, and CFI = 0.97 (>0.9). This means the items have a good fit. In addition, the t values for all items are greater than 1.96, meaning they conform with aspects of graduate competence and are valid points on which to measure them.

Result of hypothetic model testing for DIQA

The hypothetic model testing of DIQA provides evidence that DIQA accommodates the evaluation of all the input, process, and output components. The input quality influences the quality of the process, and the quality of the process influences the output.

Input Quality Evaluation: This evaluation seeks to determine the quality of an Islamic education study program by looking at (a) vision and mission of the study program; (b) the curriculum and its design; (c) the pedagogical competence of lecturers and other relevant staff, as well as professional, social, and personality competence; and (c) infrastructure and facilities,

such as mosques, classrooms, libraries, multimedia equipment, laboratories, leadership rooms, faculty rooms, administration rooms, and toilets.

Quality Process Evaluation: This evaluation seeks to assess the quality of the study program by looking at (a) the planning and implementation of processes related to learning; and (b) student coaching, such as through thesis guidance, study guidance, Al-Qur'an literacy coaching, expertise practice, and language development.

Evaluate Output Quality: This evaluation seeks to establish the quality of graduates by measuring their professionalism, such as through pedagogical competence, professional competence, social competence, and personality competence.

The results of the statistical analysis are clarified below. The DIQA test with CFA using SEM proves that DIQA has a good ability to match data (i.e., it is a good fit). Evidence from the standardized loading of the hypothetical model with component relations, variables of input quality evaluation, process quality, and output quality show that the correlational indicators among variables have a high loading factor ≥ 0.3 (Tabachnick and Fidell, 2007, p. 217; Harrington, 2009, p. 215). This means that the main indicator of latent construct for the DIQA model has been well-received and understood by the respondents, so the DIQA model's constructs have been well applied and are highly suitable for use.

Regarding the loading factor value, the evaluation of input to process quality has a loading factor value of 0.32 with a quadratic value of 0.1024. This means that 10.24% variance of input quality influences process quality. In addition, the evaluation of process quality to output quality has a loading value equal to 0.57 with quadratic value 0.3249, so 32.49% variance of process quality influences output quality. Thus, the evaluation of input quality affects the process quality and ultimately contributes to output quality. This result is reinforced by a t-value with a 5% cut-off (value t = 1.96).

Conclusion and Implication

In summary, this study makes three contributions: the requirements for an internal quality assurance, a process for developing DIQA, and empirical evidence to validate the final DIQA model.

The needs analysis was useful for internal quality assurance in the Islamic university. It covered the vision and mission of study programs, curricula, competency of lecturers and administrative staff, infrastructure and facilities, student coaching, the teaching/learning process,

assessment of the teaching/learning process, and graduate competency. DIQA comprises 10 different sets of questionnaire and a total of 480 items. It is based on the notion of CIPP to develop constructs, methods of evaluation, and procedures.

The development of items began at the exploration stage when performing the needs analysis and constructing the initial DIQA package. The validation of items then started with preliminary testing using a two-round Delphi method. The results of this were then used to develop the prototype into a model. The revised DIQA then had seven evaluation dimensions, 10 sets of questionnaire, and 477 items in a strong format. The model was then named DIQA (Delta Internal Quality Assurance).

Qualitative and quantitative techniques were employed in the development of DIQA. A qualitative approach was used to develop the prototype during preliminary testing and main field testing. Quantitatively speaking, statistical analysis using SPPS and ISREL was applied to provide evidence for the validity and reliability of the items and options in each questionnaire. The results clearly prove that DIQA is very good for consistently evaluating a study program and its individual items are valid. The final version of DIQA also improves on the questionnaires from AMI and BAN-PT.

This research, however, does have limitations in the form of less cooperative respondents, uncertain timing for the evaluation of study programs, the dilemma of respondents when measuring graduate competencies, and positioning DIQA with government accreditation. The less cooperative respondents made evaluation objectives that were not clear and did not fully match. Evaluation times also frequently did not conform, so some external validity may be lost. Determining graduate competence, meanwhile, is a very subjective matter, and comprehensive data about graduate competency is not fully available. Finally, DIQA could not gain government evaluation for accreditation.

The limitations therefore have implications for study programs and future research. Firstly, benchmarking is the ultimate goal of accreditation, and benchmarking through accreditation makes study programs strive to receive a good accreditation value by meeting all the indicators. Thus, an Islamic education program must follow the rules and standards set by BAN PT, although the specific peculiarities of Islamic values are not covered. Study programs can make use of DIQA to accommodate Islamic values, however, and an internal quality

evaluation for Islamic education programs helps produce professional teachers, which in turn helps them prepare for external evaluation.

Regarding future research, this study experienced a reluctance by respondents to become involved in the research and conduct evaluations in a timely fashion, thus affecting the validity of the items. This implies that DIQA contains less comprehensive attributes for accreditation, so aligning DIQA with BAN-PT will be problematic. Future research can help with this by verifying DIQA's items and improving both these and the dimensions of evaluation. Efforts to align DIQA as the initial accreditation system are also recommended if it is to be used by Islamic universities.

References

Abadie-Mendia, T. (2013). *Accountability and Quality in Higher Education: A Case Study. UNF Theses and Dissertations* 375, Jacksonville: University of North Florida.

Acaro, J.S. (1996). *Quality in Education: An Implementation Handbook*, Abingdon: Routledge.

Anderson, D; Johnson, R and Milligan, B. (2000). *Quality Assurance and Accreditation in Australian Higher Education: An assessment of Australian and international practice*, Canberra: DETYA.

BAN-PT (National Accreditation Body for Higher Education), (2002). *Guidelines for Internal Quality Assessment of Higher Education (2nd ed.)*, Jakarta: Ministry of National Education.

Borg, W.D and Gall, M.D. (2003). *Educational Research: an introduction*, New York: Longman.

Bridge, G.R; Judd, C.M. and Moock, P.R. (1973). *The Determinants of Educational Outcomes*, Massacusetts: Ballinger Pubhlising Company.

Chinta, R; Kebritchi, M and Ellias, J. (2016). "A conceptual framework for evaluating higher education institutions," *International Journal of Educational Management,* vol. 30, no. 6, pp. 989-1002.

Choiriyah, S. (2018). *Model Evaluasi Mutu Internal Program Studi Pendidikan Islam dalam Menghasilkan Guru Profesional,* Solo: Universitas Negeri Yogjakarta.

Ewell, P. (2010). "Twenty Years of Quality Assurance in Higher Education: What's Happened and What's Different?," *Quality in Higher Education,* vol. 16, no. 2, p. 173–175.

Fitri, Z.A. (2016). "Quality Assurance System between the Islamic University and the State University," *Aljamiah,* vol. 2, no. 2, pp. 208-235.

Fuller, A and Belkin, D. (2015). "The Watchdogs of College Education Rarely Bite," *Wall Street Journal,* 17 June 2015.

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.

Grobbelarr, S. (2007). *Data Gathering: Delphi Method. R&D in the National System of Innovation: A System, Dynamic Model*, Pretoria: University of Pretoria.

Haris, I. (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.

Harrington, D. (2009). Confirmatory factor analysis, Oxford: Oxford University Press.

Harvey, L. and Green, D. (1993). "Defining quality," *Assessment and Evaluation in Higher Evaluation,* vol. 18, no. 1, pp. 9-34.

Kelchen, R. (2017). *Higher Education Accreditation and Federal Government*, Wasington DC: Urban Institute.

Koslowski F. (2006). "Quality and assessment in context," *Quality Assurance in Education,* vol. 3, pp. 277-288.

Legčević, J and Hećimović, V. (2016). "Internal quality assurance at a higher education institution," *Poslovna izvrnost,* vol. 10, no. 2, pp. 75-87.

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,*" NEA (National Education Association), Washington, DC.

Patil, A.S. and Pudlowski, Z.J. (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.

Rezaee, R and Shokrpour, N. (2011). "Performance Assessment of Academic Departments: CIPP Model," *European Journal of Social Sciences,* vol. 23, no. 2, pp. 228-236.

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.

Roffe, I.M. (1998). "Conceptual problems of continuous quality improvement and innovation in higher education," *Quality Assurance in Education,* vol. 6, p. 74–82.

Rossi, P.H; Lipsey, MW. and Freeman, HE. (2004). *Evaluation: A systematic approach* (7th ed.), CA: Sage Publications.

Sallis, E. (2002). *Total quality management in education*, London: Kogan Page Limited.

Sekayi, D and Kennedy, A. (2017). "Qualitative Delphi Method: A Four Round Process with a Worked Example.," *The Qualitative Report,* vol. 22, no. 10, pp. 2755-2763.

Shafer, B.S and Coate, L.E. (1992) "Benchmarking in higher education: A tool for improving quality and reducing cost," *Business Officer,* vol. 26, no. 5, pp. 28-35.

Tabachnick, B.G. and Fidell, L.S. (2007). *Experimental designs using ANOVA*, Colombia: Thomson & Brooks/Cole.

Tam, M. (2001). *Assessing Quality in Higher Education by Examining the Effects of University Experiences on Learning Outcomes and Student Development*, Durham: University of Durham.

Utuka, G. (2012). *Quality Assurance in Higher Education: Comparative Analysis of Provision and Practices in Ghana and New Zealand*, Wellington: Victoria University of Wellington.

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

Watty, K. (2006). "Want to know about quality in higher education? Ask an academic," *Quality in Higher Education,* vol. 12, no. 3, pp. 291-301.

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

Yakubova, S. (2009). *Perception of Quality in Changing University Education in Kazakhstan, Kent: Kent State University College and Graduate School of Education*, Health, and Human Services.