Journal of Education, Teaching and Learning
Volume 1 Number 1 March 2016. Page 1-6
p-ISSN: 2477-5924 e-ISSN: 2477-4878

AN INTRODUCTION: EVALUATION OF QUALITY ASSURANCE FOR HIGHER EDUCATIONAL INSTITUTIONS USING RASCH MODEL

Andi Mursidi[1)], Soeharto[2)]

[1)] *STKIP Singkawang, Singkawang, Indonesia*
E-mail: andimursidi@hotmail.com

[2)]*STKIP Singkawang, Jl. STKIP Kelurahan Naram, Singkawang, Indonesia*
E-mail: soehartofisika@gmail.com

Abstract. This is a descriptive qualitatif research about quality assurance evaluation. the research aims to introduce analyzing using Rasch model to evaluate higher education institution based on quality assurance standars that have been developed to evaluate each member including instructor and staff in higher education institution. The instrument have been developed to conduct the experiment to provide raw data sample to doing practical analyzing using Rasch model in this research. The first part of this research will explain definition of the quality assurance and Rasch model analysis. The second part of this research will show introduction analysis using Rasch model to analysis sample data. The third part of this research will show a brief summary of the result and important finding in evaluation of higher assurance. Analyzing data of evaluation quality assurance using Rasch model will help higher educational institutions to increase and develop their quality assurance to be better higher educational institution.

Keywords**:** Rasch model, quality assurance, evaluation

I. INTRODUCTION

Educational Quality assurance have important role in quality development of all systems. Quality assurance evaluation is a specific assessment for individual institutes. It doesn't provide overall information of development, giving non-referable results in conclusion of the model for instructional quality assurance. Each area conducts different operations, added that the participants in evaluation of each area possess different attributes [1]. The evaluation have to provide the applicable result, enabling acquirement of more information and development of the model for instructional quality assurance in higher educational institution.

Using quality of achieving academic excellence has always been a central value in higher education. Higher education Institutions have relied on the reputation of their faculties to attract students and scholars and to give credibility to their degree programs, their graduates, and their instructors. However, the way Quality Assurance's key components, Accreditation and Evaluation or Assessment, are defined has a great influence on its implementation and impact. Assessment is about language regarding the nature of teaching, learning, and appropriate inquiry and power regarding how higher education is organized and rewarded.

Signifying the formation of higher education Quality Assurance policies in view of the transition from elite higher education to mass higher education, was marked by influences from outside the region. There is identified five broad approaches for defining quality in higher education. These are (i) quality meaning exceptional, where quality is related to conception of excellence; (ii) quality meaning perfection, where quality has consistent and error-free attributes; (iii) quality meaning fit for purpose, where quality fulfils the perceived requirements of stakeholders; (iv) quality meaning value for money; and (v) quality meaning transformation.

A. Definition of Quality Assurance in Education

Quality is described as the totality of features and characteristics of a service that bear on its ability to satisfy stated or implied needs. Quality of higher education is a multidimensional concept, which should include all its functions and activities: teaching and academic programmers, research and scholarship, staffing, students, buildings, faculties, equipment, services the community and the academic environment [2].

It has to take the form of internal self-evaluation and external review, held openly by independent specialists, if possible with international expertise, which are vital for enhancing quality. Independent national bodies have to be established and comparative standards of quality, recognized at international level, shave to be defined. Due attention should be paid to the specific institutional, national and regional contexts in order to take into account diversity and to avoid uniformity. Quality also requires that higher education should be characterized by its international dimension: exchange of knowledge, interactive networking,

Journal of Education, Teaching and Learning
Volume 1 Number 1 March 2016. Page 1-6
p-ISSN: 2477-5924 e-ISSN: 2477-4878

mobility of teachers and students, and international research projects, while taking into account the national cultural values and circumstances [2, 3]. For any higher education institution, there are several aspects of reputation which are important [2] ;

1) It is built upon the competitive elements of quality, reliability, delivery, history and price.
2) Once a higher education institution acquires a poor reputation for quality, it takes avery long time to change it.
3) Higher education reputations can quickly become national reputations.
4) The management of the competitive weapons, such as quality, can be learned like any other skill, and used to turn round a poor reputation, in time.

The movement for evaluation, came under the strong influence of the Quality assurance movement. At the same time there is also the need for measures to evaluate the performance of the institution. Within these parameters the evaluating agencies tend to adopt a number of different approaches to monitoring quality in higher education. In general, they can all be described as forms of external conditioned by the prevailing institution. some effort of governments around the world are looking for higher education to be more responsive, including making education more relevant to social and economic needs [4],

1) Widening access to higher education,
2) Expanding numbers, usually in the face of decreasing unit cost, and
3) Ensuring comparability of provisions between institutions.

B. Data Quality Assessment

There are three steps which depict in Figure 1: planning, implementation, and assessment. in the planning phase, a systematic planning procedure is used to define criteria for determining the number, location, and timing of samples (measurements) to be collected in order to produce a result with a desired level of certainty. This information, along with the sampling methods, analytical procedures, and appropriate quality assestment, is documented in the Project Plan. Data are then collected following the Project Plan specifications in the implementation phase. in the assessment phase, the data are verified and validated to ensure that the sampling and analysis protocols specified in the Project Plan were followed, and that the measurement systems were performed in accordance with the criteria specified in the ProjectPlan. Then the statistical component of data quality assestment completes the data quality assestment by providing the evaluation needed to determine if the performance and acceptance criteria developed by the planning process were achieved [5].

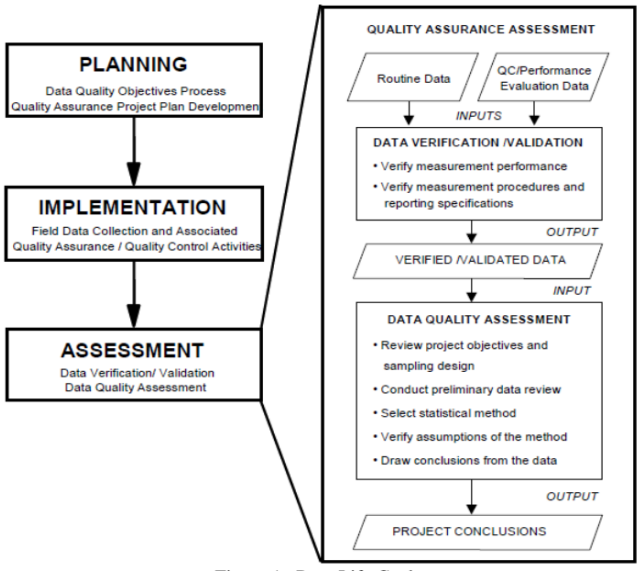

Fig. 1 Three steps in processing data quality assestment

C. Rasch Model

The Rasch model was named after the Danish mathematician Georg Rasch [6]. The model shows what should be expected in responses to items if measurement (at the metric level) is to be achieved. For the Rasch model, dichotomous [6] and polytomous [7]. The response patterns achieved are tested against what is expected, a probabilistic form of Guttman scaling [8], and a variety of fit statistics determine whether this is the case [9]. The objective is to test how well the observed data fit the expectations of the measurement model. Three overall fit statistics are considered. Two are item–person interaction statistics transformed to approximate a z score, representing a standardized normal distribution [12].

The model assumes that the probability of a given respondent affirming an item is a logistic function of the relative distance between the item location and the respondent location on a linear scale. In other words, the probability that a person will affirm an item is a logistic function of the difference between the person's level of, for example, anxiety (u) and the level of anxiety expressed by the item (b), and only a function of that difference.

$$Pni = \frac{e^{(\theta n - b\bar{i})}}{1 + e^{(\theta n - b\bar{i})}} \quad (1)$$

where Pni is the probability that person *n* will affirm the item, *u* is the person's level of anxiety, and b is the level of anxiety expressed by a positive response to the item. The formulae can be expressed as a logit model:

$$ln\left(\frac{Pni}{1 - Pni}\right) = \theta_n - b_i \quad (2)$$

where ln is the normal log, *P* is the probability of person n

Journal of Education, Teaching and Learning
Volume 1 Number 1 March 2016. Page 1-6
p-ISSN: 2477-5924 e-ISSN: 2477-4878

affirming item i; u is the person's level of anxiety, and b is the level of anxiety expressed by the item. Fitting data to the Rasch model thus places both item and person parameter estimates (note that they are independent parameters) on the same log-odds units (logit) scale, and it is this that gives the linear transformation of the raw score.

The model can be extended to the polytomous case and is known as the rating scale model [7]:

$$ln\left(\frac{Pnij}{1-Pnij}\right) = \theta_n - b_i - \tau_i \quad (3)$$

where, in addition to the parameters shown in (2) above, the t represents the threshold (0.5 probability point) between adjacent categories. A further variant of this is known as the partial credit model [10], and it makes no assumptions about the An introduction to the Rasch measurement model 3 equidistance between thresholds across items, which is the case of the rating scale model:

$$ln\left(\frac{Pnij}{1-P_{nij-1}}\right) = \theta_n - b_{ij} \quad (4)$$

Statistics indicating fit to the model test how far the observed data match that expected by the model. Note the orientation; because the model defines measurement, data are fitted to the model to see if they meet the model's expectations. This is opposite to the practice in statistical modelling where models are developed to best represent the data. Within the framework of Rasch measurement, the scale should also work in the same way, irrespective of which group (e.g. gender) is being assessed [11]. For example, in the case of measuring anxiety, males and females should have the same probability of affirming an item (in the dichotomous case), at the same level of anxiety. Thus, the probability is conditioned on the trait. If for some reason one gender did not display the same probability of affirming the item (in the dichotomous case), then this item would be deemed to display DIF, and would violate the requirement of unidimensionality [9].

A further test for unidimensionality is undertaken by looking at patterns in the residuals. These are the standardized person-item differences between the observed data and what is expected by the model for every person's response to every item. This is one way of testing the model's assumption of local independence of items; after extracting the 'Rasch factor' there should be no further pattern in the data [12].

D. Analyzing Rasch Model Using Winsteps

Figures and tables must be centered in the column. Large figures and tables may span across both columns. Any table or figure that takes up more than 1 column width must be positioned either at the top or at the bottom of the page.

Winsteps is a statistic software which develop based on Rasch model analysis. Figure 2 shows icon of winsteps with software version 3.73.

Fig 2. Winsteps icon

If we have raw data which wants to analyze using winsteps, we can drag the raw data to winstep icon. Figure 3 shows window of data setting to analyze using Rasch Model.

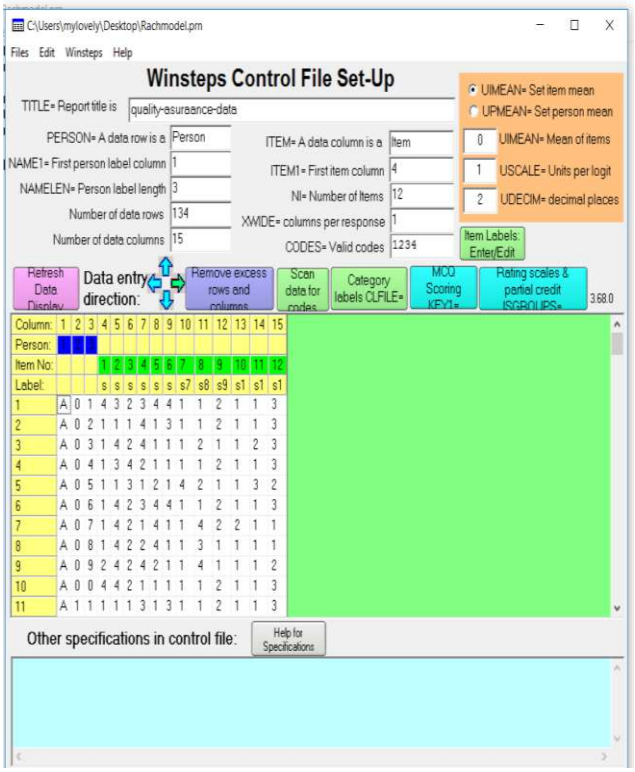

Fig 3. Winsteps window of Rasch model analyzing

After input data setting using winsteps. We can determine what is analysis which we want to present data of quality assurance evaluation. Fig. 4 shows a window of analyzing Rash model using winsteps.

Journal of Education, Teaching and Learning
Volume 1 Number 1 March 2016. Page 1-6
p-ISSN: 2477-5924 e-ISSN: 2477-4878

Output Tables | Output Files | Batch | Help | Specification | Plots | Excel/RSSST | Graphs | Data Setup

Request Subtables	1. Variable maps	20. Score table
3.2 Rating (partial credit) scale	2.2 General Keyform	21. Probability curves
2. Measure forms (all)	2.5 Category Averages	29. Empirical curves
	3.1 Summary statistics	22. Scalograms
10. Item (column): fit order	6. Person (row): fit order	7.2.1 Person Keyforms: unexpected
13. Item: measure	17. Person: measure	17.3 Person Keyforms: measure
14. Item: entry	18. Person: entry	18.3 Person Keyforms: entry
15. Item: alphabetical	19. Person: alphabetical	19.3 Person Keyforms: alphabetical
25. Item: displacement		7.2 Person Keyforms: fit order
11. Item: responses	7.1 Person: responses	
9. Item: outfit plot	5. Person: outfit plot	32. Control variable list
8. Item: infit plot	4. Person: infit plot	33. Person-Item: DGF: DIF & DPF
12. Item: map	16. Person: map	34. Comparison of two statistics
23. Item: dimensionality	24. Person: dimensionality	35. Person Paired Agreement
27. Item: subtotals	28. Person: subtotals	36. Person diagnostic PKMAPs
30. Item: DIF, between/within	31. Person: DPF, between/within	

Fig 4. Winsteps output tables

II. METHOD

This research is made using review some journal, article some literatures and analyzing data of quality assurance in higher educational institutions. Description of quality assurance of higher educational institutions have collected from some literature to defining quality assurance of higher educational institutions. Practical analyzing of Rasch model using raw sample data from 26 students to introduce analyzing evaluation of quality assurance and to introduce how to analyze data quality assurance from experiment using Rasch model.

III. DISCUSSION

According to our review evaluation of quality assurance was very importan to improve quality assurance of higher educational intitution. Quality assurance will determine reputation of higher educational institution. Some aspects which we found about reputation related quality assurance which are (1) It is built upon the competitive elements of quality, reliability, delivery, history and price. (2) Once a higher education institution acquires a poor reputation for quality, it takes a very long time to change it. (3) Higher education reputations can quickly become national reputations. And (4) the management of the competitive weapons, such as quality, can be learned like any other skill, and used to turn round a poor reputation, in time.

This research have tried to make an instrumen to measure qualit assurance using that relation with satification service of student because satification have important relation with satification service. The important one is this research trying to introduce how to evaluate data of quality assurance correctly using Rasch model analysis. Rasch model analysis have many descriptions to represent data. We use fake raw data to anayze data of quality assurance using Rasch model merely to introduce the reader how to using winsteps to use Rasch model analysis. Figure 5 and Figure 6 shows us the two of some repsentation data using Rasch model analysis.

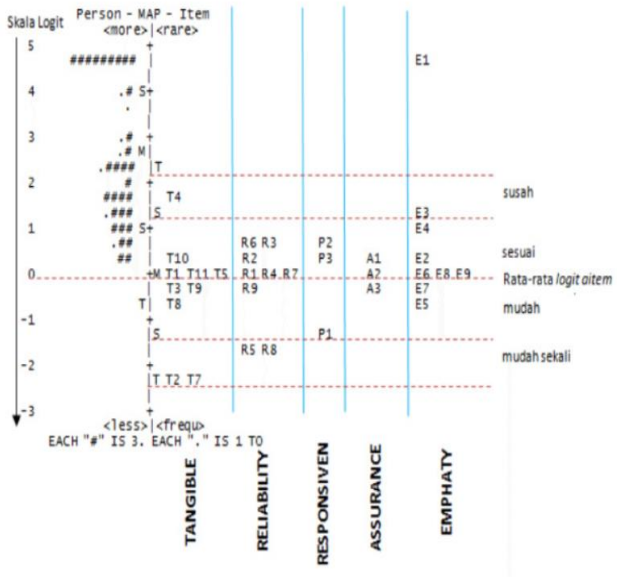

Fig 5. Wright Map with Rasch model according to category of the instrument

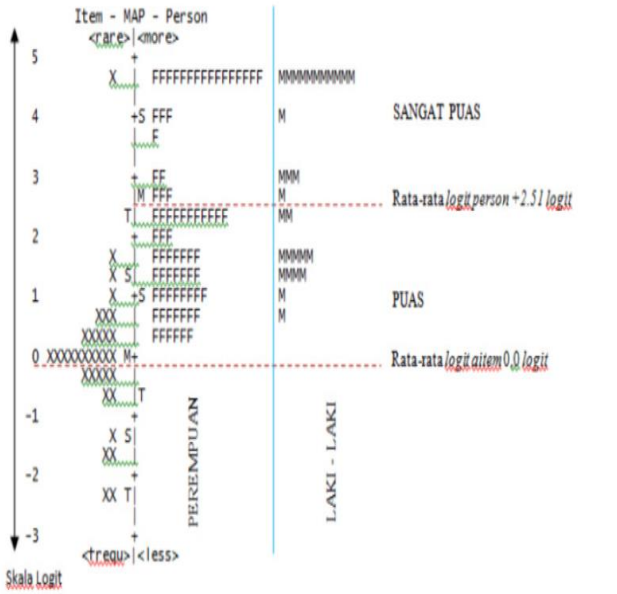

Fig 6. Wright Map with Rasch model according to satification service of female and male

Developed Instrument in this research to measure the variable quality of service. Respondent data used in mock

raw sample data. On the data generated by the instrument will rank Likert scale with five options statement. quitionary the instruments are made using the approach of five dimensions in the model SERVQUAL (Service Quality), namely Tangible, Reliability, Responsiveness, Assurance, Empathy developed by Parasuraman A, et al [13].

Analysis of testing the validity of the instrument in this study using Rasch modeling, thing seen is the validity of the response to the item based on the value outft Mean Square (MNSQ) received 0.5 <MNSQ <1.5, outft Z-Standard (ZSTD) suitability test value z received -2.0 <ZSTD <+2.0, and the Point Correlation Measure (Pt Mean corr) 0.4 <Pt Mean Corr <0.85. When a grain item instrument service quality meets at least one of the above criteria then the item worthy instrument is used, so that the instrument can be used for research.

The mock raw sample data from the measurement results in the form of data collection forms ordinal scale then transformed into an interval scale using Rasch modeling Winsteps with software version 3.73. Rasch modeling keintervalan troubleshooting data in a way to accommodate the logit transformation, by applying the logarithm of the odds ratios of raw data obtained from the respondents. The univariate analysis in this study showed levels respondents to the satisfaction of service quality dimensions seen from the distribution map of Wright item and person (characteristic) resulting from Rasch modeling.

Quality of Service Based on the analysis mock samplle data on the distribution map of Wright person map in Figure 5 it can be concluded that all respondents are above the average value of the item logit (logit +0.00) as much as 100% respondents, meaning that respondents rate the quality of service given in the category good. Based on Wright map in Figure 5 shows the left-hand column is the column of person, and the right column is kolomitem. Items that are above the average value of the item logit (logit +0.00) implies that the item is relatively difficult to be approved by the respondent and the items that were below the average logit item means the item is approved by the respondent.

Tangible column; In the column direct evidence of scattered items that easily conform until the hard-appropriate, these items have a good discrimination power than items that are in other dimensions. There is one item that is above the average value of the item logit, logit value (+1.77 logit) that the item code to the content item T4 statement "The provision of laboratories relevant to the needs of science", the item has been approved level of agreement is relatively difficult compared other items. This means the majority of respondents have inadequate laboratory facilities.

Column Reliability; On the reliability column, there are three items that are above the average value logit item with code R1 (+0.12 logit), R4 (+0.8 logit) and R7 (+0.16 logit), these items is a group of measuring the same construct, implies that these items have the same relative levels of discrimination, because the measurement results show the value logit or less the same. The third item contains statements related to the reliability of lecturers in conducting lectures.

Column Responsiveness; In the column responsiveness, there is one item under the average value logit code P1 (-1.24 logit) with the contents of the statement item "Provision lecturer Counseling for students", meaning that the item had relative approval rate almost foolproof approved significant compared to other items services have the responsiveness of the respondents.

Column Assurance; In the column guarantees there are three items to be among the average logit value item codes A1 (+0.19 logit), A2 (+0.12 logit), A3 (-0.31 logit), meaning that these items have a relatively difficult level of agreement approved by respondents. These items contains statements related to the service capability of academic staff and lecturers, meaning the quality of service is relatively more difficult to accept than the items that are in other dimensions.

Column Emphaty; In the field of attention, there is one item that is above the average value logit item (+4.71 logit) ie E1 with item code the contents of the statement item "STKIP Singkawang always tried to understand the importance and difficulties of students' level of agreement items had relatively very difficult to be approved than other items. It implies only a minority of respondents were found STKIP Singkawang always tried to understand the importance and difficulties of students.

Figure 6 shows the Quality of Service by Sex; based on the map Wright about the characteristics of the respondents indicated the left column is kolomitem, and the right column is a column of respondents. Respondents who are above the average value logit person (+2.51 logit) implies that respondents are more satisfied with the quality of services provided, compared to respondents who were below the average logit person. Characteristics by Sex; Based on the distribution map of Wright Figure 3 shows that the percentage of highly satisfied with the quality of service based group sex in order are the male sex by 55.2% while female gender was 33.8%. The result of analysis of variance between the level of assessment is very satisfied and satisfied with the quality of service based on the characteristics of gender signifkan there is a difference, the value of $F = 3.603$, $p = 0.061$ ($p > 0.05$).

Quality of service can be interpreted simply how good the level of service given to meet the expectations and needs of students are being targeted. Good service quality greatly affects the quality assurance of higher education institution. Service quality is strongly influenced by its ability to consistently meet the expectations of students.

IV. Conclusions

Quality is described as the totality of features and characteristics of a service that bear on its ability to satisfy stated or implied needs. Quality assurance has good relation with satisfication or quality service. Simply, to know about quality assurance we have to assest student satisfication toward quality service of higher educational institution. This study aimed to make good instrumen to assest quality assurance of higher educational institution and introduce how to evaluate it using Rash model analysis. We have developed instrument which result ordinal data that can

analyze using Rash model. The evaluation instrument of higher educational institution consist of five aspect, tangibles, reliability, responsiveness, assurance, emphaty. Rasch model analysis was used for analyzing student satisfication toward service of higher educational institution to evaluate quality assurance. In this paper we found that Rasch model is easy to use and have many interpretation to present and give meaning of data. Using rasch model we can now where is the point of the instrument which needs improving to increase quality assurance of higher educational institution. But finally, this study just to beginning to introduce how to evaluate quality assurance of higher educational institution, how to make instrument of quality assurance of higher educational institution, and how to analyze it using Rasch model analysis.

ACKNOWLEDGMENT

We would like to thank ADRI which provides template to help all participants writes a journal easily and conducts this international seminar. We would like to thank to STKIP Singkawang which fund all of cost to present this research and some lecturers who give advices to make better idea in this research. Without all of people who support this research, we cannot write this research well and properly.

REFERENCES

[1] Straw, R.B. and Herrell, J.M. (2002). A Framework for Understanding and Improving Multi – Site Evaluation. *New Directions for Evaluation*, 94: 5 – 15.
[2] ESIB, The National Unions of Students of Europe, "European Student Handbook on Quality Assurance in Higher Education", 2002.
[3] Ferreira, "Chapter4: Quality Models in the Higher Education Sector", University of Pretoria etd, Vol 14 NO 2 2000:182, 2003.
[4] G.Srikanthan, "Developing a Model for Quality in Higher Education", Centre for Management Quality Research, 2002.
[5] EPA "Data Quality Assessment: A Reviewer's Guide", EPA QA/G-9R,United States,Environmental Protection Agency, Office of Environmental Information, Washington, DC 20460, February 2006.
[6] Rasch, G. (1960). Probabilistic models for some intelligence and attainment tests. Chicago: University of Chicago Press.
[7] Andrich, D. (1978). Rating formulation for ordered response categories. Psychometrika, 43, 561–573.
[8] Guttman, L. A. (1950). The basis for Scalogram analysis. In S. A. Stouffer, L. A. Guttman, F. A. Suchman, P. F. Lazarsfeld, S. A. Star, & J. A. Clausen (Eds.), Studies in social psychology in World War II: Vol 4. Measurement and prediction (pp. 60–90). Princeton: Princeton University Press
[9] Smith, R. M. (2000). Fit analysis in latent trait measurement models. Journal of Applied Measurement, 2, 199–218.
[10] Masters, G. (1982). A Rasch model for partial credit scoring. Psychometrika, 47, 149–174.
[11] Holland, P. W., & Wainer, H. (1993). Differential item functioning. Hilldale, NJ: Lawrence Erlbaum Associates
[12] Julie F. Pallant and Alan Tennant. (2007). An introduction to the Rasch measurement model: An example using the Hospital Anxiety and Depression Scale (HADS). British Journal of Clinical Psychology, 46, 1–18.
[13] Kotler P, Keller KL. Marketing management. Edisi ke-14. New Jersey: Pearson Education Inc; 2012