

Quality Assurance System on Scoring Guidelines: An Example from Mathematics Department in Indonesian Distant Learning

Tutisiana Silawati*, Hasoloan Siregar

Faculty of Science and Technology, Indonesia Open University, Indonesia

Received December 11, 2019; Revised February 5, 2020; Accepted February 18, 2020

Copyright©2020 by authors, all rights reserved. Authors agree that this article remains permanently open access under the terms of the Creative Commons Attribution License 4.0 International License

Abstract The purpose of this study is to analyze the Scoring Guidelines in the assessment of students' essay test answer book of the Mathematics Department in Indonesian distant learning college. This study has been conducted in Jakarta, Yogyakarta, Surabaya, Medan, and Palembang. Students' test answer books consisted of three subjects, namely 1) the comprehensive test of MATA4500/TAP, 2) the main course of MATA4111/Calculus II, and 3) MATA4213/Numerical Methods. The sample consists of 293 students who registered during the first semester of 2014, the second semester of 2014, and the first semester of 2015. The students test answer books for each course is evaluated by two lecturers from a local college, the qualification of the lecturers are Master on Mathematic. The lecturer assessed students' essay test answer books by using a scoring guide/marketing schemes that created by the test developer from the Mathematics Departement. Data for this study were analyzed using SPSS, t-test. The results show that the differences scores of 280 out of 293 (95,55%) given by the two lecturers show that it does not exceed 5%, it is following the *Sistem Jaminan Kualitas* (Quality Assurance System) Universitas Terbuka 2013, No. Document JKOP_UJ03-PK04 PK Examination of Test Results Description. Which in turn provides that the Scoring Guidelines are good and clear.

Keywords Students' Test Answer Book, Mathematic Department, Scoring Guidelines, Quality Assurance System

1. Introduction

Universitas Terbuka (UT), is a State University that implements a distance learning system, which means students learn using media, both print, and non-print media.

Therefore UT students are expected to learn independently, by utilizing teaching materials that have been prepared for self-study by UT.

Quality Assurance, in general, refers to a process of defining and fulfilling a set of quality standards consistently and continuously to satisfy all consumers, producers, and other stakeholders. UT formulated the university's Quality Assurance System or *Sistem Jaminan Kualitas* (Simintas) in October 2001 (Tian & Amin, 2010)

Evaluation of learning outcomes implemented in the form of final exams, comprehensive tests with scientific works, and Online Tutorial. The comprehensive test is intended to verify the mastery of students in a comprehensive field of science in the program Strata one (S1). The Comprehensive test material covers some of the support courses available in the Mathematics Department. Since the first semester of 2015, the Comprehensive test score combined with scientific works that can be uploaded on UT web application via <http://karil.ut.ac.id> (*Katalog Universitas Terbuka 2017/2018. Departemen Pendidikan Nasional Universitas Terbuka, n.d.*)

A final exam may be an objective test (multiple choice) or a description test. The final exam answer is written for the objective test done on the Test Answer Sheet, where the student fills out the Test Answer Sheet using a 2B pencil. The final exam answers to the test description including comprehensive tests are done on the students' test answer book.

Assessment of students' essay test answers book conducted at five Regional namely: 1. Medan (consist of Medan, Banda Aceh, Padang, and Pekanbaru), 2. Palembang (consist of Palembang, Jambi, and Bengkulu), 3. Jakarta (consist of Jakarta, Bandar Lampung, Serang, Bogor, Bandung, Pangkal Pinang, Pontianak, Palangkaraya, Batam, and overseas), 4. Yogyakarta (consist of Yogyakarta, Surakarta, Semarang, and Purwokerto), and 5. Surabaya (consist of Surabaya,

Malang, Jember, Denpasar, Mataram, Kupang, Banjarmasin, Samarinda, Makasar, Kendari, Palu, Manado, Gorontalo, Ambon, Ternate, Majene, and Jayapura). Examination of assessment of students' test answer book involves lecturers from local universities at regional selected based on the compliance with the examiner lecturer competency courses, examination capacity (the number of lecturers examiner vs. number of students), and ease of access of Regional (Simintas, 2013).

The final exams and comprehensive test equipped with Scoring Guidelines developed by the lecturer in the Mathematics Department as an assessment guide for examiners from local universities at regional. The Scoring Guidelines along with final exams and comprehensive tests are sent by the UT-Testing Center to five regional ones day after the implementation of the final exam is completed. The results of the research are used to find out information about final exams and comprehensive tests Scoring Guidelines according to the examiners at five regional. Perfection improvements can be made which in the end the research objective to obtain a good Scoring Guideline can be done.

The Essay Test, in general, the number of test items less than the objective test, the student in answering the essay test requires a long sentence and paragraph to answer the test questions. There are two types of essay test, first explanatory test that is open answer test that is students are not given limitation in writing and organizing their answers and second limited answer test ie students are given specific limits and context in terms of form and reach of answers ("Suwanto Mengungkap Karakteristik Tes Uraian", 2010).

Students in the essay test can demonstrate their ability to interpret facts and concepts and compose answers in sequence and integrated, and the process can be measured (Writing Essay Test Items). Students' distance learning performance can be measured by evaluating learning outcomes using the essay test (Nakayama, Yamamoto, & Santiago, 2010).

Characteristics of the essay test are, (1) in general one item contains some questions; (2) answers are given in written form; (3) answers are given in long descriptions. The positive aspects of the essay test are, (1) the essay test can assess students' understanding at a high level; (2) students can present their ideas and thoughts; (3) the essay test can be prepared in a relatively short time. The lack of a description test is, (1) the value may be different when judged by the same corrector at different times and / or by different corrector at the same time; (2) the length and complexity of the answers can cause problems in the assessment; (3) the time needed to correct the problem. ("Suwanto Mengungkap Karakteristik Tes Uraian", 2010)

Assessment in the essay test has two types Analysis Method and Global Method. The Analysis Method uses an ideal answer whose value is arranged in detail step by step and each stage is assigned the value of an answer. The

value given by the lecturer is the basis of the number of stages of the answer. While the Global Method is the ideal answer is not shared in the specific stages, lecturers examine globally by assigning an overall grade of student answers ("Suwanto Mengungkap Karakteristik Tes Uraian", 2010). UT examines the results of the examination test at Regional office, the examiner is not allowed to bring home students' essay test answer book along with other examination files. Examination of test results is done by two independent examiners using the Scoring Guidelines. The maximum limit of value difference between the two examiners, for non-exact courses is 10%.

As for the exact course is 5%. If the difference in value between the two examiners exceeds the maximum limit, then the two examiners re-examine the students' test answer book. (Simintas, 2013)

The Mathematics Department is supported by students through the Final Semester Examination (UAS), there are two types of first evaluations, the multiple-choice examinations, and the written examinations. Written examinations are developed by lecturers in the UT Mathematics Department (Tutisiana & Zulmahdi, 2015), while the assessment is carried out by lecturers or examiners from other universities who have a mathematical background with master and doctor of philosophy qualifications located at 5 Regional namely: Jakarta, Yogyakarta, Surabaya, Medan, and Palembang. Examination of student exam results is carried out by 2 examiners to maintain the objectivity of the assessment of student exam results maintained. To reduce subjectivity a Test Examination Guideline is needed for examiners. As with the multiple-choice test, the student's answer is checked by the machine (Personal communication with Evaluation experts, DR. Herman 02/05/2019). Furthermore, Bonnie L. MacGregor (MacGregor, n.d.) states that clearly-made Examination Guidelines are an important tool for viewing performance or assessing student exam results. This opinion is also supported by Tutisiana and Arter (Tutisiana & Zulmahdi, 2015) (Arter, 2010) stating that specially designed Examination Guidelines can provide clear information about student examination results.

UT implements learning using a distance learning system. As a consequence, assessment of students' essay answers book conducted at five Regional, so that scoring guidelines are the necessary guidance for examiners. The scoring guide is the lecturer's reference in scoring the student's answers that contain the possible answers and scores appropriate for each possible answer (Simintas, 2013).

The Scoring Guidelines come with a score for each answer process or stages of answered made by the student. Guidelines were developed as a guideline Assessment scoring for lecturers who checked students' essay test answers book. Therefore, lecturers at five Regional examining students' essay test answers book with the same standard. Every one student essay test answers book

examined by 2 correctors, the grade came from the average between the first corrector and the second corrector.

2. Methods

The study was conducted at UT's Examination Centers at March-April 2016, intended to study how much different grades of students between the corrector/examiner 1 with 2. Regarding the students' essay test answers book which has been examined. The three courses examined namely: 1) MATA4500 / TAP, is the comprehensive test, 2) MATA4111 / Calculus II, is of competence subject, and 3) MATA4213 / Numerical Method, this course is registered by students representing 5 Regional. Students' essay test answers book examined for 3 exams periods: first semester of 2014, the second semester of 2014, and the first semester of 2015. The survey data for this study were collected from the students' essay test answer books for three courses at 3 exam periods, as shown in Table 1. Data for this study were analyzed using SPSS, t-test.

Table 1. Courses and Students' Essay Test Answer Books

No	Code/Course	Students' Essay Test Answers Books
1	MATA4500/Comprehensive	20
2	MATA4111/Calculus II	160
3	MATA4213/Numerical Method	113
Total		293

Table 1 stated that there were 20 Students' Essay Test Answers Books for Comprehensive Course, 160 Calculus, and 113 Numerical Method, thus the total of Students' Essay Test Answer Books was 293.

3. Results

The t-Test results for all courses between the 1st examiner and the 2nd examiner, shown that no difference. All this can be seen from sig (2-tailed) in an independent sample test in the t-test column. All values are significant at $p > 0.05$, which means there is no difference between the assessments given by the two examiners. Thus two examiners from the local colleges give the same value for the students for the three Courses namely 1) the comprehensive test of MATA4500/TAP, 2) the main course of MATA4111/Calculus II, and 3) a course taken by the students in five regions of MATA4213/Numerical Methods.

3.1. Assessment by Examiners at Regional Centers for MATA4111/Calculus II

The t-Test results in for MATA411/ Calculus II courses. All this can be seen from sig (2-tailed) in an independent

sample test in the t-test column. All values are significant at $p > 0.05$, which means there is no difference between the assessments given by the two examiners, shown that no difference between the 1st examiner and the 2nd examiner. Thus the two examiners from local colleges give the same judgment for MATA4111/Calculus II Students' Essay Test Answer Books.

3.2. Assessment by Examiners at Regional for MATA4213/Numerical Methods

The t-Test results for MATA4213/Numerical Methods courses shown that no difference between the 1st examiner and the 2nd examiner. All this can be seen from sig (2-tailed) in an independent sample test in the t-test column. All values are significant at $p > 0.05$. Stated that the two local colleges give the same value for the Students' Essay Test Answer Books for MATA4213/Numerical Methods.

3.3. Assessment by Examiners at Regional for MATA4500/Comprehensive Test

The t-Test results for MATA4500/Comprehensive Test courses shown no difference between the two examiners. All this can be seen from sig (2-tailed) in an independent sample test in the t-test column. All values are significant at $p > 0.05$, which means there is no difference between the assessments given by the 1st examiner and the 2nd examiner. As a result, no difference between the two examiners in gives judgments for the Students' Essay Test Answer Books for MATA4500/Comprehensive Test.

4. Discussion

The t-Test results for either the course or for each course indicate no difference between the 1st examiner and the 2nd examiner. All this can be seen from sig (2-tailed) in an independent sample test in the t-test column. All values are significant at $p > 0.05$, which means there is no difference between the assessments given by the examiner 1 and 2.

Based on the results of the study, the Scoring Guidelines, which were developed by the lecturer at the Mathematics Department, which was used as a guide by two university lecturers in the Regional to assess the results of the student test, successfully managed to guide the examiners to examine the test results with the same perception. Scoring Guidelines can guide two lecturers well, thus there is no difference in assessment between the examiner one and the second examiner, on the student's written exam results. This means that the Scoring Guidelines are valid, that is, they can carry out their measurement functions appropriately and reliably, that is, shows consistency as a measuring tool and ensures the objectivity of the assessment.

Besides that, according to Ron Legion (Ron, 2015). Scoring Guidelines have the advantage of being able to

improve communication between academic institutions in improving the quality of distance education. Furthermore, Scoring Guidelines is an evaluation instrument that can be used to evaluate an assessment, because the Scoring Guidelines can assure the accuracy of the assessment and can minimize differences in perception and subjectivity of the lectures (I Made S.U.).

The advantages of the Scoring Guidelines are:

1. Reduce the time spent grading by following the instruction from Scoring Guidelines.
2. Facilitation examiners ensure consistency across time and grading.
3. Decrease differences in perception and subjectivity element in conducting an assessment.

5. Conclusions

The finding of this research is as follow. A benefit of using a Scoring Guidelines for the examiner is to have the criteria and the scoring standards defined to facilitatng the activities of correcting the students' essay test answer book. Besides, the two examiners stated that they have the same perception for the Scoring Guidelines. With the result that, the examiner gives with the same perception, thus there is no difference between the assessments given by the 1st examiner and the 2nd examiner. In my opinion, according to this research, the Scoring Guidelines are worthwhile. Scoring Guidelines can reduce the subjectivity of evaluators to test results, so students are not disservices. Scoring Guidelines can be applied to other universities that use the distance education system.

REFERENCES

- [1] Arter, J. A. Scoring Rubrics. In *International Encyclopedia of Education* (pp. 123–139). Retrieved from <https://doi.org/10.1016/b978-0-08-044894-7.00262-1>.
- [2] IMadeS.U. Panduan Latihan Membuat Rubrik. Retrieved from <http://repo.unand.ac.id/3526/46/14.%20PANDUAN-LATIHAN-PEMBUATAN-RUBRIK.pdf>
- [3] Katalog Universitas Terbuka 2017/2018. Departemen pendidikan Nasional UniversitasTerbuka. (n.d.).
- [4] MacGregor, B. L. (n.d.). Rubrics as an Assessment Tool in Distance Education. In *Encyclopedia of Distance Learning* (pp. 1583–1588)A. Abiewskiro, Z. Moplskiiera. The Problem Of Grammar Choice For Verification, TCSET of the International Conference, House of Lviv Polytechnic National University, 19-23, 2008.
- [5] Ron. (2016). Legon "Measuring the Impact of the quality matters rubrikTM: A Discussion of Possibilities. *American Journal of Distance Education*, 29 (2015). Retrieved from <https://www.tandfonline.com/doi/full/10.1080/08923647.2015.1058114>.
- [6] Simintas, P. (2013) JKOP UJ03 Pemeriksaan Ujian Retrieved from <http://pusmintas.ut.ac.id/index.php/pedoman-simintas/pedoman-simintas>
- [7] Suwarto (2010). Mengungkap Karakteristik Tes Uraian. *Jurnal Widyatama*.
- [8] Tian, B., & Amin, Z. (2010). "The Practice of a Quality Assurance System in Open and Distance Learning: A case study at Universitas Terbuka. In *of Research in Open and Distance Learning* (Vol. 8, pp. 173–178). Retrieved from <http://www.irrodl.org/index.php/irrodl/article/view/340/774>
- [9] Tutisiana, S., & Zulmahdi, D. (2015). Hasoloan S, "Penilaian Korektor Terhadap Pedoman Penskoran Ujian Uraian Matakuliah-matakuliah yang diasuh oleh PS Matematika."
- [10] Rubrics: useful assessment tools. Centre for Teaching Excellence, University of Waterloo Retrieved from <https://uwaterloo.ca/centre-for-teaching-excellence/teaching-resources/teaching-tips/assessing-student-work/grading-and-feedback/rubrics-useful-assessment-tools>