Application of Learning Management Test Instruments To Identify Questions As Evaluation Tools

Abdul Hayat¹, Nur Azizah², Amalia Nur Rizqi³
¹ Institut Teknologi Bandung, Jl. Ganesha No.10, Lb. Siliwangi, Bandung, Jawa Barat
² Universitas Gunadarma, Jl Margonda Raya Nomor. 100, Depok, Jawa Barat
³ STMIK Raharja, Jl. Jendral Sudirman No.40 Modernland, Cikokol, Tangerang
e-mail: abdul.hayat@raharja.info, nur.azizah@raharja.info, amalia.rizqi@raharja.info

Abstract

In the era of disruptive 4.0, the process of education in management at Higher Education was held in various forms such as formal education and non-formal education. In addition, the application of learning management as an evaluation tool also aims to find and find various factors that lead to the success and failure of students after following the education process, so that a bright spot will be found as a way out for improvement in the implementation of education in the future. This study discusses the notion of tests, the development of tests as an evaluation tool, true-false tests, multiple choice tests, various forms of multiple choice tests, matching tests, matching tests (completion test), and short answer questions. Therefore, it is necessary to conduct an analysis of the instrument for making the items that have been made by management at the University in evaluating their students. With the hope that the question analysis aims to identify the questions that are good, not good, and questions that are not feasible. Given to students.

Keywords: Test Instruments, Evaluation Tools, and Management

1. Introduction

In the world of education where the learning process in learning management is highly supported by technology to determine the quality of graduates of a university. Therefore at this time to achieve the objectives in learning management apply the test as an evaluation tool with the aim of knowing and measuring the success of the learning management process in Higher Education [1]. The test is a set of questions that must be answered, must be addressed, or tasks that must be carried out by students who take the test. Tests are used to measure the extent to which a student has mastered the lessons delivered, especially covering aspects of knowledge and skills so as to facilitate management of learning in assessing student with abilities [2].

Tests on learning management are generally intended to measure aspects of human behavior, such as aspects of knowledge (cognitive), aspects of attitude (affective), and aspects of skills (psychomotor). The thing that is to be measured is the level of mastery of students on the learning material that has been taught. In this description, it is necessary to distinguish between learning achievement and learning outcomes. Learning achievement only measures two aspects, namely cognitive and psychomotor aspects. The results of learning management include aspects of the character formation of a student, thus measuring the three main aspects of educational outcomes, namely cognitive, psychomotor, and affective aspects [3].
Table 1. Examples of Questionnaire Instruments and Tests

<table>
<thead>
<tr>
<th>Variabel</th>
<th>Sub Variabel</th>
<th>Indikator</th>
<th>Item Soal</th>
<th>Banyaknya Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hasil Belajar Matematika</td>
<td></td>
<td>Kognitif</td>
<td></td>
<td>1,2,3,4,5,6,7,8,9,10</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Afektif</td>
<td>Hasil Belajar dalam cara: Intelektual, Penubahan Sikap</td>
<td>11,12</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Psikomotor (Sadjana, 2005)</td>
<td>Hasil Belajar dalam cara: Keterampilan</td>
<td>12,14,15</td>
</tr>
</tbody>
</table>

In terms of learning management using this test instrument we distinguish between two forms of tests, namely as follows: First, subjective tests, which are generally in the form of essays. According to Ganjar Nugraha, M. (2018), essay form tests are a kind of learning progress test that requires answers that are discussion or description of words [4]. The characteristics of the question are preceded by words such as; describe, explain, why, how, compare, conclude, etc. [5]. There are not many questions about essay forms, only about 5 - 10 questions in about 90 to 120 minutes. These essay form questions require the ability of students to be able to organize, interpret, connect the meanings that have been owned based on the management of Education that has been set. In short, it can be said that essay tests require students to be able to remember and re-recognize, and especially must have high creativity so that they can improve the effectiveness of research in education management [6].

According to Larasati, S. (2018), the goodness of subjective tests on learning management can be explained in 4 (four) points, namely (a) easily prepared and compiled, (b) does not give many opportunities to speculate or chancy, (c) Encourage students to dare to express opinions and arrange in good sentences. Give the opportunity to students to express their intentions in their own language style and ways, and (d) to know the extent to which students explore a problem being tested [7].

An objective test is a test that can be carried out objectively. This is indeed intended to overcome the weaknesses of essay form tests so that learning management conducts research on test instruments using multiple choices for student evaluation tools. In the use of this objective test the number of questions put forward is far more than the essay test [8]. Sometimes tests that last for 60 minutes can be given 30-40 questions. The virtues of objective tests, namely: (a) contain more positive aspects, for example more representative represent the content and extent of material, more objective, can be avoided the interference of subjective elements both in terms of students and the aspects of the examining teacher; (b) it is easier and faster how to check it because it can use test keys and even the results of technological advancements; (c) the examination can be submitted by another person; and (d) in examinations, there is no subjective element that affects [9].

2. Research Method

In carrying out the design and development of a system there are several very important things that must be considered, including the existence of research methods to solve various formulation of the problems that exist in the stages of this research.
In this study using several research methods including observation techniques, and Literature Review. In the observation technique has been carried out directly by monitoring the multiple-choice test as an evaluation tool in a learning management at the University.

A. Literature Review

According to Handayani, I., (2018), at the time of this management research researchers took place using the literature review method which had a definition of a method used to gather relevant information in accordance with the topics and problems that were the object of research [10].

According to Rahardja, U. (2019), Literature Review is written material both in the form of books and sources from previous research as management of learning, discussing the issue of journals or scientific works with topics to be researched that are relevant to previous research topics or those that already exist. Because in previous studies there have been many scientific studies that conducted research on credit memo features to minimize deposits and bad debts [11].

1. Research conducted by Aini, Q., Rahardja, U., Moeins, A., and Apriani, D. M. (2018), entitled "Application of Gamification in Student Exam Assessment Information System to Improve Lecturer Performance". This study reviews a system where the development of gamification concepts in the PEN + (Plus Assessment) system can be effective because lecturers input the UTS, UAS, and Independent Tasks in a timely manner so that learning management becomes more effective and efficient. The second objective is the victory trophy for 10 (ten) gamification ranking lecturers who are appreciated by lecturers who have entered the value on time so that they can motivate lecturers to compete in the process of entering values on time so that the quality of lecturers can be well formed and facilitate management in decision making. With the development of gamification in the PEN + (Plus Assessment) system in learning management in Higher Education can improve the quality of lecturers and students [12].

2. The research was conducted by Amrullah, A., (2018), entitled "Analysis of the School Examination Multiple Choice Question Test Instrument (US) for Al Islam Class XII Muhammadiyah High School throughout Palembang". This study reviews a learning management study which results in 3 (three) conclusions, namely the first item validity multiple choice exam school Al-Islam subject class XII second semester Muhammadiyah high school throughout Palembang city 2015/2016 academic year is valid. And the second result is the reliability of the multiple choice questions in the Al-Islam school class XII Semester 2 in Palembang city in the 2015/2016 academic year included in the test which has reliability that is equal to 0.914, which is much higher than the standard coefficient of 0.700. The moderate level of difficulty of the 50 questions presented is 4 items or about 8% items. And in the final study the tests included in the category of difficult questions consisting of 36 items or 72% included in the category of questions that are currently consisting of 10 items so the percentage of as much as 20% in learning management so that it is more effective and efficient [13].

3. Research conducted by Sari, D. R. U., Wahyuni, S., and Bachtiar, R. W., (2018), entitled "Development of Instruments for Multiple Choice High Order Thinking Tests on E-Learning-Based Physics Learning in High School". This study reviews the management of learning using a test instrument where expert validation data meets valid criteria and the empirical validity obtained meets very high validity criteria. The results of the reliability test on the product show that the analysis of the items, namely the level of difficulty of packages A and B, still strongly requires improvement based on education management. In the analysis of differentiation there are 27 questions worthy of use and 13 questions in categories not feasible because they have a bad distinguishing power. In the analysis of deception there are 16 questions package A and 10 questions package B with distractors accepted because it is quite good and there are 14 questions with distractors revised because it is not good and (3) students' high level
thinking ability after using test instruments Multiple choice high order thinking in physics e-learning based on high school can be categorized quite well [14].

4. The research was conducted by Pratiwi, H. Y. (2017), with the title "Development of Multiple Choice Test Instruments to Identify the Characteristics of the Thermodynamic Concept of Students in Malang Kanjuruan University Physics Education Study Program". This study reviews the results developed in the form of a multiple choice distractor test instrument to identify the characteristics of student physics conceptions that have been successfully carried out in accordance with the management's requirements in Higher Education. The number of questions in the test instrument developed was 30 items, which were developed from 10 indicators. The results of the analysis stated that most of the items of the multiple choice test developed were declared feasible in terms of material, construction, and language. With the hope that this research can develop assessment instruments as an effort to identify the characteristics of the thermodynamic concept with the instruments used are multiple choice tests on education management. Data collection was done through questionnaires, interviews, description tests, and multiple choice tests of meaningful distractors [15].

5. Research conducted by Tiara, K., Kusumah, H., and Putri, D. M., (2017), entitled "Application of Asset Management with Mappyourtag at Higher Education". This study reviews a system with barcodes where the asset management process that exists at Raharja College runs effectively without having to print the checklist form and the data is automatically updated in a database system so that management decisions on managed assets can have effective functions that is by using the MapYourTag application. Where in the previous system asset management applied was still manual or very conventional, namely by writing reports manually, then inputting the report into a data processing application. The results of this study are in the form of an asset management application system by scanning barcodes that have been created for each asset that is connected directly to the database that has a dashboard that shows the statistics of control treatment on each asset, so that supervision becomes more effective, optimal and easier in making asset management reports. And with the existence of this system, it is expected that asset management implemented at Raharja College can improve sustainable development in the asset life cycle [16].

Based on the 5 (five) literature review on the research method, it is explained that there is an important role of learning management in providing a multiple choice test instrument as an evaluation tool to identify good, poor, and inappropriate questions. nowadays there is still a great need for learning and understanding of the preparation of test instruments in order to evaluate the success and failure of students. So the researchers conducted research and development of a compilation of test instruments to identify questions with case studies relevant to management where there was no topic in previous studies.

3. Results and Analysis

The preparation and development of tests is intended to obtain valid tests, so that the results of the measurements can reflect precisely the learning outcomes or learning achievements achieved by each individual test participant after participating in learning management activities [17]. For that the steps of the construction of the tests taken are as follows:

1. Set test objectives. Learning achievement tests can be made for various purposes, such as: (a) tests that aim to hold national examinations or other similar tests; (b) tests that aim to hold a selection, for example SNMPTN; and (c) tests that aim to diagnose student learning difficulties known as diagnostic tests.

2. Curriculum analysis. Curriculum analysis aims to determine the weight of each subject that will be used as the basis for determining the number of items or items for each objective.

Application of Learning Management Test Instruments To Identify Questions As Evaluation Tools (Abdul Hayat)
subject matter or the weight of the questions for the description form, in making the test grid.

3. Analysis of textbooks and other learning material sources. This analysis has the same purpose as curriculum analysis, which is to determine the weight of each subject. But in textbook analysis determine the weight of each subject based on the number of pages of material contained in the textbook or other learning material sources.

4. Make a grid. The benefits of grids are to guarantee samples of good questions, in the sense of covering all the subjects proportionally. In order for the items or items of the test to cover the entire material (subject matter or sub-topic) proportionally, then before writing the test items we must first make a grid as a guide.

5. Writing Specific Instructional Objectives (ICT). Writing ICT must be in accordance with the stipulated conditions. ICT must reflect student behavior, therefore it must be formulated operationally, and technically use operational words.

6. Writing questions. After the grid in the form of a specification table is available, then we will make the items in question. Some instructions that need to be considered in making items (especially mathematics for example) are:
   a. The questions made must be valid (construct validity) in the sense that they are able to measure the achievement of the learning objectives that have been formulated.
   b. Questions made must be able to be done using one specific ability, without being influenced by other irrelevant abilities.
   c. The questions made must first be done or completed with complete steps before being used on the actual test.
   d. Establish the beginning of the aspect of ability to be measured for each math problem made.
   e. In making mathematical questions, avoid as far as possible typos no matter how small, because it will affect the validity of the question.
   f. Provide instructions for working on the questions completely and clearly for each form of mathematical questions in a test.

7. Question review (face validity). The questions that are made are still possible shortages or errors involving aspects of specific abilities measured, language used, typos and so on. For that before being reproduced, the questions must first be examined by colleagues who understand the test material and the question writing technique to examine the surface validity of the questions made.

8. Limited test reproduction. Prepared tests are reproduced in sufficient quantities according to the number of trial samples or the number of participants who will take the test in a test trial activity.

9. Test trials. Tests that have been multiplied will be tested on a number of predetermined samples. The trial sample must have characteristics that are more or less the same as the characteristics of the actual test taker.

10. Test trials. Tests that have been multiplied will be tested on a number of predetermined samples. The trial sample must have characteristics that are more or less the same as the characteristics of the actual test taker.

11. Question revision. Valid questions based on the empirical validity criteria are confirmed by the grid. If the questions meet the requirements and have represented all the material to be tested, the questions are then assembled into a test, but if valid questions have not met the requirements based on the confirmation results with the grid, improvements can be made to the problem required.

12. Assembling questions becomes a test. The order of the questions in a test is carried out according to the level of difficulty of the question, which is from easy questions to difficult questions.

Judging from the purpose and function the test is divided into 4 (four), namely: (1) placement test, (2) formative test, (3) diagnostic test, and (4) summative test [18]. Formative
tests are tests conducted at the end of each discussion of a subject or topic, and are intended to determine the extent to which a learning process has proceeded as planned [19].

Summative test is a test conducted at the end of each unit of time which includes more than one subject, and is intended to determine the extent to which students have been able to move from one unit to the next [20]. According to Purnanto, AW, and Suryawan, A., (2018), defining summative evaluation as the use of tests at the end of a particular teaching period, which includes several or all units of learning taught in one semester, even after discussion of a field of study [21].

Diagnostic evaluation is an evaluation that is used to determine the strengths and weaknesses that exist in students so that they can be given the right treatment. Diagnostic evaluation can be carried out in several stages, both at the initial stage, during the process, and at the end of learning. In the initial stage, it is done to prospective students as input while in the final stages of this diagnostic evaluation to determine the level of student mastery over all the material that has been studied [22].

**Table 2. Comparison of Diagnostic Tests, Formative Tests, and Summative Tests**

<table>
<thead>
<tr>
<th></th>
<th>Diagnostic Test</th>
<th>Formative Test</th>
<th>Summative Test</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Function</strong></td>
<td>mengelompokkan siswa berdasarkan kemampuannya menentukan kesulitan belajar yang dialami</td>
<td>Umpan balik bagi siswa, guru maupun program untuk menilai pelaksanaan suatu unit program</td>
<td>Membantu tanda telah mengikuti suatu program, dan menentukan posisi kemampuan siswa dibandingkan dengan anggota kelompoknya</td>
</tr>
<tr>
<td><strong>Evaluation Method</strong></td>
<td>Memilih tiap-tiap keterampilan prasarata. Memilih tujuan setiap program pembelajaran secara berimbang. Memilih yang berhubungan dengan tingkah laku fisik, mental dan perasaan</td>
<td>Mengukur semua tujuan instruksional khusus</td>
<td>Mengukur tujuan instruksional umum</td>
</tr>
<tr>
<td><strong>Scoring (Order)</strong></td>
<td>Menggunakan standar mutlak dan relatif</td>
<td>Menggunakan standar mutlak</td>
<td>Menggunakan standar relatif</td>
</tr>
</tbody>
</table>

1. **Test is True-False (True-False)**

   The right-wrong form is 2 (two) types which are seen in terms of working or answering questions, namely (1) with corrections (with correction) that is the student is asked to correct if he chooses the wrong answer, and (2) without correction (without correction) ie students are only asked to circle the letter B or S without giving the correct answer. Test goodness is wrong, is (1) can cover a wide range of materials and does not take up much space because usually the questions are short, (2) easy to compile, (3) can be used multiple times, (4) can be seen quickly and objective, and (5) instructions on how to do it are easy to understand. While the weaknesses are (a) often confusing, (b) easily predictable or predictable, (c) many problems that cannot be expressed only with two possibilities right or wrong and (d) can only reveal memory and re-recognition [23].
2. **Multiple Choice Test**

According to Farida, A., (2018), many levels and levels of education use multiple choice tests. This is because: (a) the type of test is compiled and used to measure all competency standards, from the simplest to the most complex (b) the number of alternative answers (options) is more than two so that it can reduce guessing (guessing) (c) this type of test requires the ability of students to distinguish various levels of truth at once and (d) the level of difficulty of the item can be controlled by simply changing the level of homogeneity of alternative answers. Multiple choice formative tests are scored objectively, because the examination or scoring is not always done by humans but can be done by machines such as scanner machines. Multiple choice test consists of a statement or notification about an incomplete understanding. And to complete it must choose one of several possible answers that have been provided. Or multiple choice test consists of a description (stem) and a possible answer or alternative (options) section. Possible answers (options) consist of one correct answer namely answer key and some distractor [24].

3. **Ordinary Multiple Choice Test**

According to Diana, E., and Hasanuddin, H., (2018), a multiple choice test can be used to measure more complex memory skills, comprehension, and application. This form of test can also be used to measure a student's higher ability and can be secretly objective. The usual multiple choice test consists of the main sentence in the form of an incomplete statement. To complete the perfection of the sentence the application of answer choices must be in the form of the answer chosen to complete the statement. The incomplete statement in the form of this question is indicated by the presence of vacancies or points that need to be filled in to complete the statement [25].

4. **Association Multiple Choice Test**

According to Handayani, I., and Dewanto, I.J., (2018), there are several analytical techniques that have been proposed by experts in this field, among others, are deceitful. This is according to Brown's opinion as quoted by that the choice of answers can be eliminated or revised unless someone chooses it. With that basis, it is known the effectiveness of deceiving the answer choices. Test makers find it difficult to develop test items that are able to measure the level of understanding rather than directly measuring knowledge of teaching material. Variations and ways to compile objective test items to measure achievement of teaching goals are often found in more complex forms of testing. Answer choices such as "all the answers above", "none of the answers above", "one of the following", all the answer choices like that can make the test participant have difficulty. In addition, making all choices of right or wrong answers allows the test taker to choose the best or closest to the actual answer choice [26].

This form of association multiple choice type is almost the same as the usual multiple choice type, the difference is that more than one correct answer is possible. For the form of association multiple choice tests, basically it is almost the same as the usual multiple choice form, but in this form the answer method is more complex [27]. Examples of items are as follows: for my item, choose:

A. If only (1), (2), and (3) are correct
B. If only (1) and (3) are correct
C. If only (2) and (4) are correct
D. If only (4) is correct.
E. If everything is correct

5. **Matching Test**

According to Nasafi, I., (2018), explain the advantages and disadvantages of matchmaking questions. Its advantages are: (a) the breadth of material that can be covered (b) questions are relatively easier to make, especially when compared to multiple choice questions
concise and economical in terms of how to provide answers and (d) can be done easily, quickly, and easily in calculating the score. While the weaknesses are: (a) there is a tendency to measure memory ability and not properly used to measure higher cognitive abilities; (b) the difficulty of maintaining homogeneity in the contents of the premise and response, especially in terms of the similarity of abilities to be measured; and (c) the possibility of guessing correctly is relatively high, because the number of question statements (in the left hand lane) with the answer statement (in the right lane) is not much different.

Example: "Pair the questions in the left lane with those in the right lane by placing the letters in front of the left lane statement at the points provided in the right lane."

- c. Transmigration .......................... 1. Entry of residents from other countries.
- a. Immigration.............................. 2. Movements of residents to other countries.
- d. Emigration ............................... 3. Moving residents from village to city.
- ............................................... 4. Movements of residents between islands in one country.

4. Conclusion

From the results of the research and discussion that has been done, it can be concluded that the test is a set of questions that must be answered, must be addressed, or tasks that must be carried out by students who take the test. The preparation and development of tests is intended to obtain valid tests, so that the results of the measurements can reflect precisely the learning outcomes or learning achievements achieved by each individual test participant after participating in teaching and learning activities making it easier for management to make a decision. Judging from the purpose and function the test is divided into four, namely: (a) placement test; (b) formative tests; (c) diagnostic tests; and (d) summative tests. Objective tests that are often used in learning evaluations are: (a) true-false tests; (b) multiple choice tests: (a) true-false tests; (b) multiple choice tests: (c) the matter of matching; and (d) filling test.

References