



Electronic Testing as a Tool for Optimizing the Process of Control over the Results of Educational Training Activities

Zhanna V. Chaikina^(✉), Sofya M. Shevchenko, Maria V. Mukhina, Olga V. Katkova, and Lyubov I. Kutepova

Nizhny Novgorod State Pedagogical University Named After Kozma Minin (Minin University), Nizhny Novgorod, Russian Federation
jannachaykina@mail.ru, shevchenko.sm@mail.ru,
lubovkutepova@mail.ru, mariyamuhina@yandex.ru,
katkova.ov@yandex.ru

Abstract. The article deals with the problem of ensuring the objectivity of the pedagogical control of results of the academic activity of students. One of the leading modern means of pedagogical control in educational institutions of higher education is pedagogical testing, carried out with the help of information and communication technologies. In K. Minin Nizhny Novgorod State Pedagogical University for the implementation of effective educational activities in the main areas of implementation and training profiles for students, the electronic educational environment LMS Moodle has been used. The authors of the article consider the electronic educational environment of LMS Moodle to be an effective tool for creating control and measuring materials for assessing the results of the learning activity of students. The article reveals the possibilities of this electronic environment for the creation by teachers of the university of electronic educational materials, means of control, including pedagogical tests. The article describes the algorithm for creating parallel forms of electronic tests based on the tool capabilities of the electronic resource of the LMS Moodle environment. The authors of the article suggest using “random questions” to create parallel (multi-variant) forms of tests. “Random questions” are added to the test directly at the design stage and regulate the inclusion of a specific task from the e-course issues bank in this question. In addition, the authors give the results of an experimental study of the effectiveness of the use of electronic tests created in the electronic environment of LMS Moodle based on parallel forms of the test.

Keywords: Pedagogical testing · Pedagogical control
Electronic educational environment · Electronic tests
Parallel forms of tests · Test tasks

1 Introduction

At present, large-scale optimization of the work of higher educational institutions of Russia is being carried out. A well-developed information and communication technology’s (ICT) infrastructure is being formed, and a massive increase in the

qualification of higher school employees is in parallel. Teachers of higher educational institutions more and more successfully use electronic educational resources in the educational process.

According to the latest regulatory documents of the education system, computer (electronic) testing is defined as pedagogical testing using a computer running a special program that is designed to provide the required presentation of test questions and the necessary processing of test results.

The importance of electronic testing in the learning process can not be overestimated. The main advantages of electronic testing include: The ability to simulate test tasks; Efficiency of evaluation; Objectivity of evaluations; Possibility of reflection, feedback from students; Ease of use [3, 4].

2 Theoretical Basis of the Research

The problem of designing pedagogical tests is considered in numerous studies of domestic [1, 9] and foreign authors [2, 4, 10]. The electronic educational environment of LMS Moodle includes a rich arsenal of various tools for the creation of test, control and measurement materials for evaluating the learning outcomes of students. The main element in LMS Moodle for testing knowledge of students is a test that allows the teacher to ask different question formats and ways of selecting answers.

The environment of LMS Moodle has a great potential for creating test tasks of different types, types and forms of presentation. This electronic environment makes it possible within the course to create a test and to set up a system for evaluating the answers of students for certain learning tasks. In order to fully realize these opportunities, the teacher needs to devote a lot of time to the process of designing assignments.

On the available tools of the LMS Moodle environment, the teacher can prepare a bank of test tasks according to his electronic course [3, 8].

Using these tools, the teacher - the author of the electronic course can make a test task of any complexity, focusing on the level of student preparation and specific training tasks: From simple tests containing popular and original types of answers, to complex test tasks that combine several questions.

When creating test tasks in the LMS Moodle environment, it is required to distinguish the functions of such elements of the electronic environment as the "Test Tasks Bank" and "Test". Function the questions bank cumulates all the test tasks of this course, allows the teacher to systematize and manage certain volumes of test questions of the electronic course on discipline, and provides access to test tasks from published categories of questions of other author's courses.

The test is an element of the electronic course, with which the learner directly works, he contains a certain set of test tasks collected in a certain order.

In modern studies, the term "Pedagogical test" is considered as a representative pedagogical system of test tasks of increasing difficulty, a specific form, which allows qualitatively evaluate the structure and measure the level of preparedness of subjects [5].

Tests used to evaluate the results of training activities should have several options [1]. Since, this makes it difficult to write-off and guess during the control event, which makes it easier to conduct simultaneous testing of large groups of students. In addition,

one of the test quality indicators is the reliability of test results, which in practice is usually considered as a test characteristic reflecting accuracy of test measurements, and the stability of test results to the action of random factors [9].

The problem of determining the reliability of tests and test tasks is considered in the studies of Russian [1, 9] and foreign authors [4]. Reliability is often defined as a measure of the correlation between two parallel test cases. Therefore, ideally, we should try to make several variants of the test, and for this, different versions of the same task must be developed.

When developing parallel forms of tests, experts recommend using a facet, the so-called special form of recording several variants of the content of the same test task. This principle is fundamental when creating parallel tests.

Theoretically, before control activities with the help of the developed test, it is necessary to test for a separate sample of students with a minimum of 100 people in order to reject poorly performing tasks. This stage (it can be termed as a “test approbation”) is mandatory in the preparation of test tasks and the construction of tests. However, in the conditions of the educational process, most often the empirical testing and testing of the test is carried out during the training on the first group of students.

The correctness of each pre-compiled test task (question) can definitely be indicated only by the results of an empirical study. Only after carrying out an experimental check of the test results and their statistical processing is the conclusion about the validity and reliability of the test results. Domestic experts note that in this case it is necessary to talk about the validity and reliability of the test results, and not the test itself [1]. Transformation of tasks in the test form into test tasks begins with the moment of statistical check of each task for the presence of test-forming properties. Each task, before becoming a test, must go through a stage of statistical and test case study.

The electronic educational environment LMS Moodle has a certain set of tools for assessing the quality of developed tests. Element of the course “Statistics” allows the teacher to easily determine the effectiveness of the created test tasks. This procedure becomes possible only after the trial testing of students using the created test. The process of empirical evaluation of test quality and test tasks is described in sufficient detail in the domestic scientific literature [5].

The main indicators of the effectiveness of the developed tests in the electronic environment of LMS Moodle are standard deviation, the index of differentiation and the coefficient of differentiation.

- (1) Standard deviation: The indicator reflecting the scatter of points that the students received when answering a certain test question. In the event that all students respond to the assignment in the same way, the scores will be zero. This result indicates that this task is not a test task and, therefore, should be excluded from the test.
- (2) The index of differentiation: Is a coefficient of discrimination, that is, an indicator of the ability of each specific task to separate the best students from the worst. This index takes values between +1 and -1. In the event that a differentiation index with a negative value is obtained, it means that the subjects from the strong group answer this question better than the strong ones. Such test tasks should be discarded, since they actually reduce the accuracy of the entire testing procedure.

- 3) The coefficient of differentiation: The second way to measure the ability of a particular test task is to differentiate between strong and weak subjects. This parameter can also take values between +1 and -1. This coefficient shows the relationship between the results of the students performing the test as a whole and the answers received by the subjects in the performance of a specific task.

In the event that the values of this coefficient have positive values, it means that the used test tasks do differentiate the trainees with high and low level of preparation. When the coefficient takes negative values, this circumstance indicates that poorly trained students respond to these test tasks better than well-prepared ones. Similar tasks with a negative value of the differentiation coefficient can not be called test, since they do not meet the basic requirements of testing tasks related to the assessment of the level of student preparation. Such tasks should be excluded from the test.

The coefficient of differentiation is more sensitive for detecting the effectiveness of the measuring ability of test questions in comparison with the index of differentiation. The advantage is that the former uses the data of the whole set of learners, and not just the results of more successful and less successful students.

The pedagogical test is a system of tasks assembled in a certain way, each of which is necessary for performing a test of its function. Removing at least one job from the test system results in a space in the list of tested knowledge, which reduces the quality of measurements. To increase the accuracy of the measurement, the test uses the location of the test tasks (questions) in order of increasing complexity. The test includes the minimum number of tasks necessary to obtain accurate results.

3 Methodology of the Study

The purpose of the study was to theoretically substantiate and develop a technology for assessing the quality of training of university students using multivariate (parallel) forms of tests and experimentally test the effectiveness of implementation.

Two electronic courses developed in the electronic environment of Moodle, existing on the basis of the University of Minnesota, participated in pilot test: "Cutting of materials", "Fundamentals of technical creativity". Electronic courses in these disciplines have been used in e-learning practice for more than two years. In addition, test tasks from the banks of questions of these courses were preliminarily tested and adjusted.

For the experiment, two parallel tests were developed for each course: Experimental and control on certain topics of the course with the same number of tasks and the same difficulty for the intermediate control of students' knowledge. The tasks in the test were selected open and closed: The question with a short answer, the question of multiple choice, the question of correspondence, the question of restoring the sequence, the tasks of the experimental and control tests were not duplicated, but in form of presentation, content and difficulty were equivalent. The questions in the tests were arranged according to the degree of difficulty increase, the function of mixing questions in the test was turned off.

These tests contained 30 questions. The time for electronic testing was laid in the program and was 30 min.

Each question was rated at 1 point; therefore, the maximum possible score is 30, which corresponds to 100% of the performance.

At the end of the time provided for testing, LMS Moodle automatically completes the testing process and displays the number of points and the final result in the form of % completion of the submitted tasks on the monitor screen. The student had the right to finish testing before the expiration of the given time.

The control tests for each subject were the same and had no options. The construction of experimental tests for each discipline had goal creation of a multivariate means of control, which ensures the objectivity and reliability of the results of the evaluation.

In order to prepare multivariate (parallel) tests with several presentation options for the test subjects, the LMS Moodle program was used to develop the tests, namely, creating a test task in the “body” of the test in the form of a random question. These random questions do not have their own content, they do not contain the text of the assignment and the answers. Random questions added to the test directly at the stage of constructing the test, perform a “programming” function that regulates the inclusion of a specific task from the e-course issues bank in this question. To ensure that the function “add a random question” is included in the content of the test. Questions are created in the questions bank in which categories the questions of a certain content and form are cumulated, to which the program applies while creating a random question.

In the course of the research, an algorithm for creating a multivariate (parallel) electronic test was developed, which includes the following steps:

- (1) The allocation of a module containing the subjects of the discipline subject to control;
- (2) The content of each topic is divided into separate educational elements - didactic units. Didactic units are the minimum units of discipline content, which are usually oriented to one or another educational task. This educational problem can be connected with science (phenomenon, information, hypothesis, proof, theorem, axiom, experiment, fact, knowledge, theory, principle, method of research, etc.) or practice (skill and an indicative basis of activity, etc.);
- (3) The specification of the test is developed, the number of test tasks for each training element and the test as a whole (the length of the test) is determined;
- (4) Each test task (initial), according to the principle of facetedness, is created in several variants;
- (5) In the questions bank, categories are created according to the number of test tasks (initial), these categories can be combined into larger categories according to the developed educational elements of the module topics;
- (6) Test tasks created by the principle of facetedness are stored in categories (subcategories);
- (7) The test is constructed by the way of adding a random question to the “body” of the test; when creating the question, a subcategory with a set of facet test tasks is indicated;
- (8) The test settings for electronic testing are carried out.

The values of the standard deviation, index and coefficient of differentiation were taken into account when assessing the effectiveness of test tasks (questions) in the trial testing. In the trial (preliminary) testing, mainly students studying in the field of training “Pedagogical education” and “Operation of transport-technological machines and complexes”, who work in the considered electronic courses of disciplines, took part.

In the experimental and control tests, only those test tasks were included, according to which the index values and the differentiation coefficient were not less than 0.3. Test tasks with lower or negative values of the index and differentiation coefficient, as well as assignments with the value of the mean square deviation close to zero, were rejected and not included in final versions of the tests.

4 Analysis of the Results of the Study

When processing the test results for control and experimental tests of e-courses, the indicators were used, which were determined using the electronic system LMS Moodle, namely the average score and the median estimate of the milestones (see Table 1).

Table 1. Results of testing based on the use of control and experimental tests

Criteria	OTT Course		Course “RM”	
	CT	ET	CT	ET
Average rating from all attempts (%)	72.38	55.07	76.42	59.31
Median score for all attempts (%)	76.67	58.00	79.05	60.87
Scale of points	18.3	23.1	16.2	20.8

The study showed that the distribution of the results is not symmetrical for both control and experimental tests, the average value of individual scores deviates to the right from the middle of the scores, counting from 1 to 30, which indicates a certain ease of the tasks of the tests being analyzed. In addition, the test tasks of the control tests were easier for the students.

At the same time, the average score based on testing results using experimental tests of all e-learning courses is closer to 50% in value than in control tests. This circumstance testifies to the reliability of the test measurements [9].

The difference between the maximum value of the primary score and minimum value is called the scores of the subjects. Other things being equal, the best test is the one whose scores are higher. In this connection, the results of testing with the help of experimental tests are more reliable than the results of control tests compiled in different electronic courses of disciplines.

Thus, the study showed that in the process of monitoring the results of training activities of students it is more effective to use electronic tests created on the basis of random questions, which increases the objectivity of control and reduces the probability of copying. The electronic educational environment LMS Moodle has sufficient

capabilities for the development of multivariate (parallel) forms of tests that facilitate the implementation of control measures to assess the learning achievements of students.

References

1. Avanesov, B.C.: Composition of Test Tasks. Training Book, 3 edn., 240 pp. Testing Center (2002)
2. Greaney, V., Kellagan, T.: Assessment of Educational Achievements at the National Level, 108 pp. T.N. Leonova, Scientific, Logos (2011). Chelyshkova, M.B. (ed.)
3. Gruzdeva, M.L.: The method of information modeling as a means of learning and an instrument of knowledge of reality. *Bull. Univ. Minin* **2**(10), 13 (2015)
4. Crocker, L., Algina, J., Naidenova, N.N., Simkina, V.N., Chelyshkova, M.B.: Introduction to the classical and modern theory of tests. In: Zvonnikova, V.I., Chelyshkova, M.B. (eds.) *Under the Society*, 668 pp. Logos (2010)
5. Krasnov, Y.E.: Guidelines for the development of test tasks and the construction of pedagogical tests. <http://charko.narod.ru/tekst/metodiki/krasnov.pdf>. Accessed 20 Mar 2017
6. Terekhina, O.S., Chaikina, Z.V.: Informational educational environment in the organization of independent work in the university. In: *The Collection: Integration of Information Technologies into the System of Additional Education in the Field of Technical Creativity, A Collection of Articles on the Materials of a Regional Scientific and Practical Conference*, pp. 37–40. Nizhny Novgorod State Pedagogical University. K. Minin (2016)
7. Terekhina, O.S., Chaikina, Z.V.: Information and communication technologies as a means of activating the independent activity of students - future technology teachers. *World of Science*. vol. 4(3), p. 22 (2016)
8. Chaikina, Z.V., Smirnova, Z.V.: Use of information and communication technologies in pedagogical testing. *Mod. High Technol.* **11-2**, 392–396 (2016)
9. Chelyshkova, M.B.: *Theory and Practice of Designing Pedagogical Tests*, p. 432. Logos, Moscow (2002)
10. Roid, G.H., Haladyna, T.M.: *Technology for Test-Item Writing*. Academic Press, New York (2008)