

Optimization of electronic test parameters in learning management systems

Yevhen Palamarchuk^a, Olena Kovalenko^a

^aVinnitsia National Technical University, Khmenytske ave, 95, 21021, Vinnitsia, Ukraine

Abstract

The article contains the results of research on the procedures for creating and adjusting tests in learning management systems based on an active dialogue with students and automated analysis of test results. The authors performed research on the procedures for creating, evaluating quality, and adjusting tests based on the electronic learning management system JetIQ.

To evaluate the optimized method of creating and adjusting tests in learning management systems, the authors use modules to assess the quality of tests, feedback; modules for the analysis of answers to questions. Students assessed by tests do this several times: by topic; on the intermediate control of knowledge (colloquium) and the final control of knowledge (exam).

This approach allows to make changes to the procedure of the final assessment of the knowledge, make adjustments to questions, and select the most correct to combine in the exam. The resulting student activity profile allows the teacher to be more objective when using test scores. The results of the evaluation of the method of optimization of test adjustment procedures are also presented, which indicate a significant effect in saving time.

Keywords

learning management information system, knowledge testing module, quality feedback module, answer analysis module, "smart test", JetIQ VNTU

1. Introduction

The learning management system should be an information ecosystem and cover all educational processes. The principle of the ecosystem involves the reuse of information that is entered once into the system. Among the various modules of the learning management system on the example of the author's system, JetIQ VNTU can be distinguished primarily by the office of teacher and student.

They form the basis of the information ecosystem. Educational processes are also automated using the electronic dean's office module. Information is provided through news systems and jet sites of departments. The office of the teacher of the learning management system contains various modules. The "My repository" module is used to download electronic resources. Providing access to students under the educational programs of the specialty is implemented using the module Navigator of educational resources of the discipline. The IQ test module is used to control knowledge by testing students.

ICT&ES-2020: Information-Communication Technologies & Embedded Systems, November 12, 2020, Mykolaiv, Ukraine

EMAIL: p@vntu.edu.ua (Y. Palamarchuk); ok@vntu.edu.ua (O. Kovalenko)

URL: <https://vntu.edu.ua/> (Y. Palamarchuk); <https://vntu.edu.ua/> (O. Kovalenko)

ORCID: 0000-0002-7443-099X (Y. Palamarchuk); 0000-0001-7116-9338 (O. Kovalenko)

© 2020 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)

Improving testing modules in the learning management system is a topical issue for educational institutions that actively use testing tools. The format of distance and blended learning involves the use of tests both during training and during control activities.

Among the known practical approaches to the formation of test tasks and improvement of testing tools in learning management systems include the use of tools for quantitative and qualitative assessment of students' knowledge in learning management systems, analysis of learning outcomes for the formation of test groups [1]. Among them: tests and quizzes, exercises, written works, individual interviews, special activity reports [2].

The testing module is one of the most complex and intelligent modules for the following reasons:

- The testing module should provide the opportunity to create questions of different types and conduct testing in different modes - training and examination.
- The testing module should make it possible to assess the quality of the test according to certain parameters and, ideally, such an assessment should be carried out automatically.
- Test validation and verification procedures should also be automated.

In various platforms of learning management systems (Moodle, Collaborator etc.) [3, 4], as well as some test software applications, the issue of creating and combining tests is partially solved, and the procedures for their validation, verification, and quality assessment are not considered at all. That is why research to improve the procedures for creating and adjusting tests is relevant.

2. Related Work

Best practice in using tests involves using a test to assess students' input knowledge; during training to strengthen the acquired knowledge and skills. In addition, it is advisable to use questions in different tests to identify errors and features [5].

The practice of active use of feedback tests, especially instant feedback, allows to open a new level of dialogue "student-teacher", in addition, the teacher will better form an understanding of the student's knowledge and activity [6]. Research shows that the organization of feedback online or in blended learning allows to create an active learning environment, and the quality and detail of feedback, the accumulation of statistics also determine the level of both the course and learning [7].

Active use of learning management systems involves the use of various tools for assessing the knowledge and activity of the student. We can note that student activity can be effectively used both to motivate learning and to collect the use of this data to improve the quality of testing.

Distance and blended learning involves the active use of formative and final assessment of students' knowledge through tests with built-in modules for the accumulation of statistics and feedback. Such tools for analytical evaluation of the obtained results are partially implemented in different LMS, but are not used fully enough. One example is the use of analytical assessment in a civil engineering course at an Australian university [8].

Student activity, their assessment is the basis for forming a student profile. It is based not only on test statistics but also on other student activities. In particular, they can receive statuses, awards, medals for activity, evaluation of dialogues, for the number of tests performed, qualitative assessment of the knowledge, etc.

This comprehensive approach to assessing and motivating students allows them to form their profile and obtain quantitative and qualitative assessments for the study of the discipline.

Providing students with information about their achievements activates the mechanisms of motivation and increases the level of their activity in the e-learning system and includes additional feedback loops.

This allows to run procedures to actively verify the quality of test questions based on teacher-student relationships. Current training during the school year is optimal in terms of time to gather information on the quality of test questions and identify those that do not meet the requirements.

The experience of developing and implementing a training management system and supporting scientific and methodological activities of teachers JetIQ VNTU [9, 10] allows us to draw conclusions about the need for automated modules for assessing the quality of tests, automation of verification procedures and incident detection. The organization of the use of tests at the university, at the faculty, by individual teachers depends on the level of implementation of the learning management system and the policy of using tests to assess knowledge at the university [11, 1].

The general data of the student's activity include completed tests, tasks, files sent to teachers, participation in lectures, use of lecture material, dialogues in chat and forum, etc. (Fig. 1).

3. Proposed Methods and Materials

The well-known method of trial and error is actively used by teachers around the world when creating tests. The main phases of this method are presented in Fig. 1. The test will be called a pool of questions on a particular topic and/or the entire discipline. This approach has its historical roots in the use of paper-based testing. But even today the phase of improving the test, which is created in an electronic system that has no special modules for quality assessment and verification of issues.

Development. The duration of this phase depends on many parameters:

1. The total number of questions in the tests;
2. Areas of knowledge;
3. Availability of formulated closed and open answers;
4. Availability of calculation parts in the test questions;
5. Availability of graphics in questions and answers;
6. Use of tests in which several correct answers are chosen;
7. Using tests to compare questions and answers.

It is extremely difficult to identify the main factor influencing the time of test development and therefore we can only roughly focus on data on test development based on the experience of VNTU teachers.

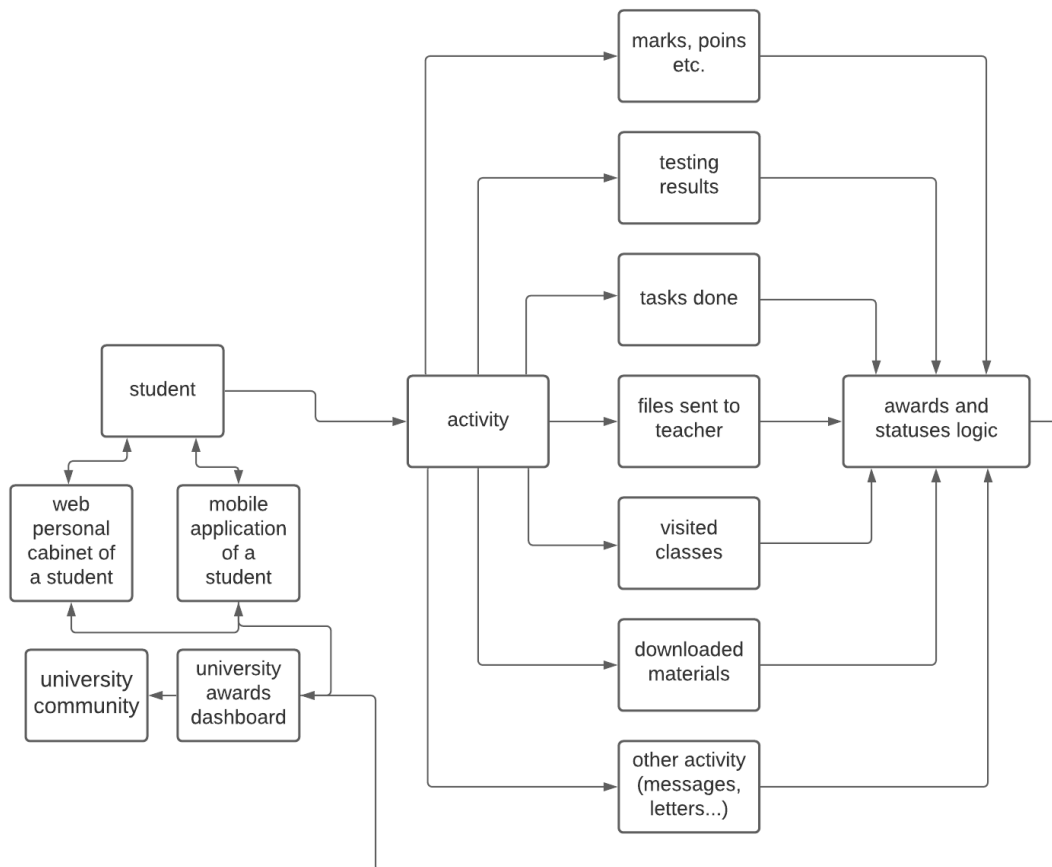


Figure 1: Scheme of the motivation of student activity in the JetIQ system

Table 1 presents approximate data of test development time taking into account the above parameters.

The testing process can be carried out periodically according to the schedule of the educational process. For example, it can be separate control measures, check of separate subjects of discipline, etc. Also, testing can be conducted continuously, provided that there are tests on each topic or laboratory and practical classes.

The third phase involves verifying the test results. Its duration depends on the number of incidents and the number and types of questions in the tests. If the testing is carried out in a learning management system that has a special incident detection module, the duration of this phase can be ignored.

The main data for correcting the tests are students' complaints about unclear questions or incorrect answers included in the program. Markers for correcting tests are also unambiguously correct and incorrect answers of students to certain questions.

Another method is to search for questionable answers in terms of the identity of the student's actual knowledge and test results. Such questions should also be corrected in the test database.

Table 1

Approximate data of test development time

Field of knowledge	Number of teachers	Total number of questions	The presence of closed answers	Availability of open answers	Availability of graphics	Estimated development time
Information Technology	2	150	3-10	5-15	2-60	10-40 days
Engineering	2	150	3-10	5-15	2-60	40-100 days
Energy	2	150	3-10	5-15	2-60	50-60 days

The optimized method of creating and adjusting the test is based on a systematic approach to working with test questions and their use in various formats. In the JetIQ learning management system, teachers can use the following tools:

1. Formation of tests with closed answers.
2. Formation of tests with open answers.
3. Using graphics in questions.
4. Using graphics in answers.
5. Forming tests with one correct answer.
6. Forming tests with many correct answers.
7. Use tests for comparison (formed as a closed answer).
8. Use tests from randomized input parameters and the calculation part. These tests do not allow students to remember correct answers and require the efforts of students to perform the calculation.
9. The use of claim politest matches for combining on a variety of topics.
10. Using tests for self-study.
11. Use exam test matches.
12. To control the results are used:
13. Quality assessment module.
14. Module for monitoring the results of answers to a separate question.
15. Test module import from a word processor.
16. Feedback module when a student finds an error.

4. Case study

Consider the features of the procedures for creating and adjusting the test using special electronic modules of the JetIQ VNTU system.

The teacher creates a test, publishes it in an electronic test system. Students take this test and receive test results. They form an information base for the automatic analysis of test data.

All the data obtained on the evaluation of the quality of tests, errors in them, the distribution of scores in the questions form an information base for the adjustment of the test.

If the tests are presented on paper and are static in any electronic testing system, or such systems do not have analytical units for assessing the quality of tests, then all the procedures

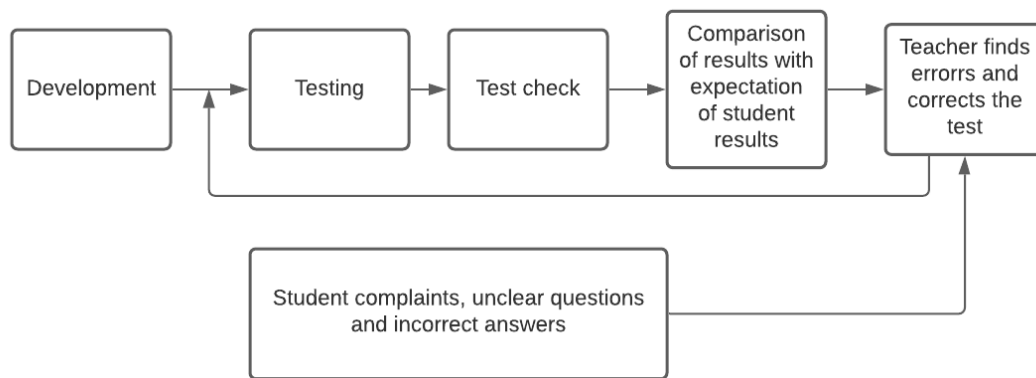


Figure 2: Scheme of student activity profile formation

for adjusting the tests are carried out by the teacher himself. After adjusting the test, it can be re-applied.

Improving the quality of tests is a repetitive procedure that is performed cyclically. Subject to the formation of tests on topics and the assessment of students several times, such an adjustment can be made according to the results of each test. The level of test quality increases gradually by implementing the following steps:

- Error questions are corrected after errors are identified by the teacher and / or students;
- The issue of the correct derived mainly or mostly wrong answers holes in them are not qualitatively formulated and well adjusted;
- To increase the level of quality assessment, the teacher should try to design the test in such a way that on the one hand, the student's answers to the questions allow to assess the level of knowledge as accurately as possible. On the other hand, test questions should be designed in such a way as to minimize the percentage of guessing the correct answers. In our opinion, this criterion is best met by questions with randomized input conditions and a computational test program. It is also important to have a large enough number of questions. With our estimates for one discipline, they should be at least 80-100 for one credit of the discipline.

The experience of using the testing unit in the JetIQ VNTU system shows that such an adjustment is widely used during the training process. Also, tests are used not only on individual topics but also on intermediate control measures. Their corrected questions are the basis for combining the final exam tests.

Also, many student activities are recorded in the e-learning system. These include completed tests, tasks sent to teachers, participation in lectures, use of lecture material, dialogues in the chat and forum, etc. (Fig. 3).

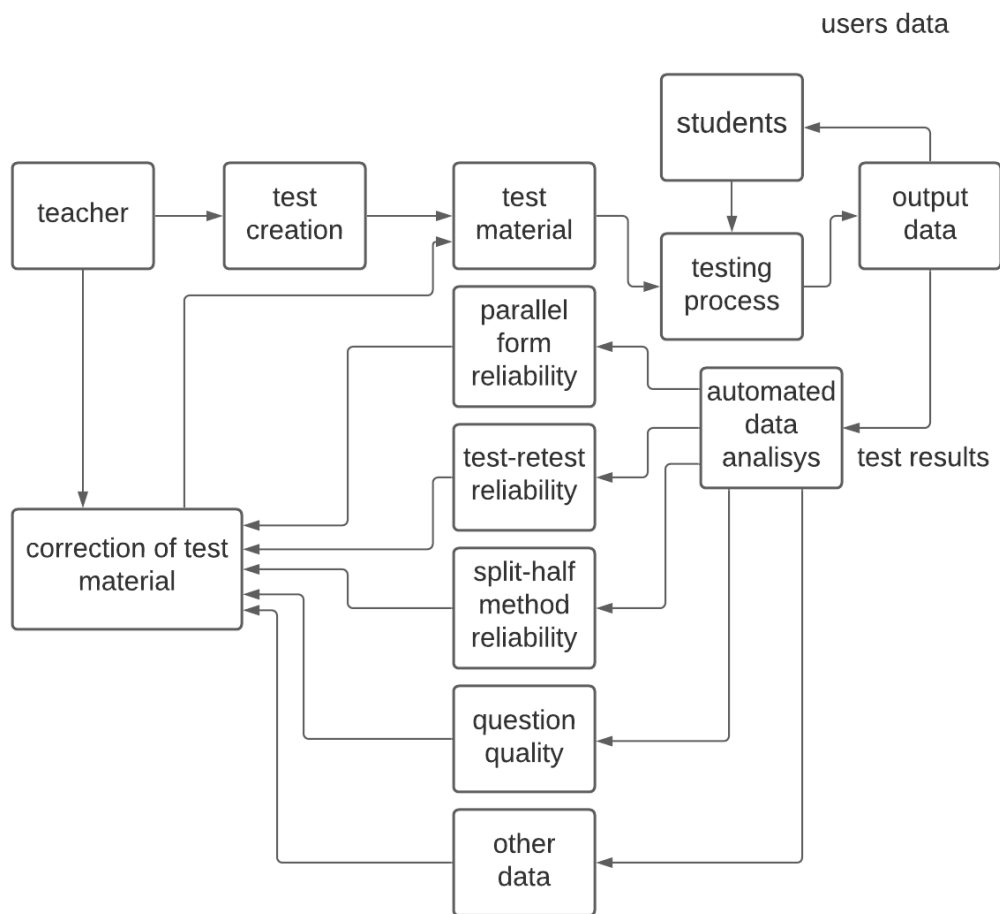


Figure 3: Diagram creation and corresponding correcting test management system with special training modules

Such work with tests is a motivation for students to increase the level of their activity in the e-learning system and gives the student a sense of partnership in correcting questions. In general, feedback on the level of correctness of questions is one of the types of dialogue with the student during testing. This comprehensive approach to assessing and motivating students allows to form a student profile, to understand how active he was during the study of the discipline.

The student's profile is based not only on statistical data of passing tests but also on other activities of the student, in particular the badges received by him, medals for activity, estimation of dialogues, a quantity of passing of tests, qualitative estimation of knowledge, etc.

Let's estimate the duration of the process of optimization of the quality of the test at the use of an electronic system of training. The duration of the stage of its creation is t_{ro} , similar to the creation of tests on any medium and in any electronic system. The main time is spent on the

formation of questions, answer options, calculation tasks, comparison tests, preparation and implementation of graphics, checking the formed tests, and reviewing - as students see them.

Let $t_{ro} - t_{d1}$, be the testing time of students. After testing, the modules of the electronic system form the results of the analysis of the quality of the test and its components ($t_{d1} - t_{s1}$).

According to the analysis, the teacher forms the necessary changes and makes the system $t_{s1} - t_{d2}$, the number of errors and inaccuracies in the questions in the previous stage was reduced, then for the next stage their number usually decreases, ie the duration of subsequent phases decreases from $t_{d2} - t_{s2}$, , which is why the organization of time for training tests should not be limited. Repeated passing of tests is recommended for students in the mode of studying materials. This will allow the teacher to make adjustments if necessary almost continuously before the exams. The results of the adjustment affect the growth of the test quality $q_0 - q_n$, and decrease the variance in the answers to questions $d_1 - d_n$.

Calculate the test optimization time. Let t_r , be the time to create the test; $t_{cn} - t_{cn-1} = t_{so} - t_{dn}$, - time to correct questions in the nth iteration. The total time to create and optimize the test is determined by:

$$T_{\Sigma} = t_r + \sum_{i=2}^n t_{cn} - t_{cn-1} \quad (1)$$

Let's perform a situational calculation of the required number of adjustments to the test questions. Let the test have Q_0 , questions that need adjustment. The test will be considered adjusted if $Q_0 < 1$,

Let's estimate the adjustment time on the n -th, phase of the iteration, taking into account the types of questions. For a test that has a total number of questions N on the n -th, iteration, the results of the correction form a general adjustment Q , which consists of adjusting the following types of questions Q_a , - a question with an ambiguous answer; Q_g , - questions of easy guessing; Q_i , - question of incorrect wording; Q_e , - questions with incorrect answers.

$$Q = Q_a + Q_g + Q_i + Q_e. \quad (2)$$

For each stage, the calculation equation for Q f will be similar (the number of questions by type will be indicated in lower case)

$$Q_f = q_a + q_g + q_i + q_e. \quad (3)$$

We will assume that in the process of their correction the teacher may make mistakes with probabilities $p_a; p_q; p_i; p_e$ according to each type of question. Then the number of correctly adjusted questions can be calculated as

$$Q_n = Q_a * (1 - p_a) + Q_q * (1 - p_q) + Q_i * (1 - p_i) + Q_e * (1 - p_e) \quad (4)$$

We introduce correction coefficients ϵ , which will characterize the ratio of questions with errors to the number of fixed for each type of question. Consider the case where the questions belong to the same type $\epsilon = 1 - p$ After the first adjustment we have the following number of incorrect questions:

$$Q_1 = Q_0 * \epsilon = Q_0 * p \quad (5)$$

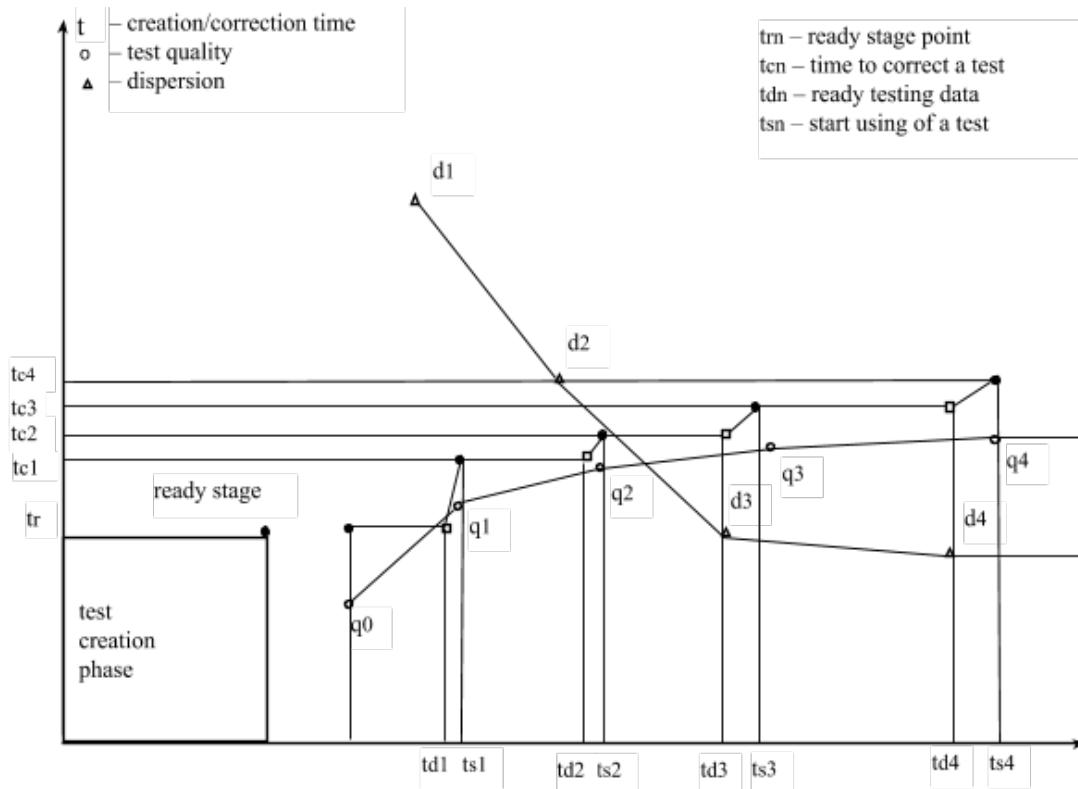


Figure 4: Phases of duration of creation and adjustment of the test

For the general case of n-repetitions

$$Q_n = Q_0 * p^n \quad (6)$$

In case of no errors

$$Q_n = Q_0 * p^n < 1 \quad (7)$$

From the last formula, we can conclude that the required number of iterations should be equal to

$$n > \frac{\ln \frac{1}{Q_0}}{\ln p} \quad (8)$$

If $p < 1$ can be written as

$$n > \frac{\ln Q_0}{\ln p} \quad (9)$$

For the presence of questions of different types of expression can be represented as follows

$$n > \frac{\ln q_a}{\ln p_a} + \frac{\ln q_g}{\ln p_g} + \frac{\ln q_i}{\ln p_i} + \frac{\ln q_e}{\ln p_e} \quad (10)$$

We introduce the value t , which will characterize the average time of the teacher to correct one question of a certain type. Then the total time to adjust the test questions can be written as

$$T = \frac{\ln q_a}{\ln p_a} \Delta t_a + \frac{\ln q_g}{\ln p_g} \Delta t_g + \frac{\ln q_i}{\ln p_i} \Delta t_i + \frac{\ln q_e}{\ln p_e} \Delta t_e \quad (11)$$

Note that the calculations are valid for cases where students pass incorrect questions in a sufficient number of times R . Therefore, the minimum duration of repeated $t_{dn} - t_{sn-1}$ (Fig. 4) should depend on this value and the time interval between the phases of the test τ . Determine the factors that affect the total number of passes R . Consider the ideal case when R reaches the minimum value of 1. Then, for the total number of questions in the test Q , when it passes S students and each of them is randomly offered r questions, we can write

$$R = \frac{Q}{S * r} \quad (12)$$

To reliably diagnose incorrect questions, the number of tests done should be as large as possible. But in practice, this number is limited to a certain amount of M .

$$R = \frac{Q * M}{S * r} \quad (13)$$

The duration of the interval $t_{dn} - t_{sn-1}$

$$t_{dn} - t_{sn-1} = R * (\tau_i + \tau_t) = \frac{Q * M}{S * r} * (\tau_i + \tau_t) \quad (14)$$

Where the interval τ_i - characterizes the period between the possibility of re-passing the test and τ_s - the average time to answer the test questions.

For example, exams and control tests can take several months. This significantly complicates obtaining the necessary amount of data for quality diagnosis of incorrect issues. For training tests, the interval τ_i can be significantly lower. In the case of automated training systems, it can be reduced to zero. Then the minimum value of the phase of the period between the possibilities of passing the retest τ_i is reduced to zero (no special passage is required, this is done by students in the phase of training and / or intermediate control and the duration of the interval $t_{dn}t_{sn-1}$

$$t_{dn} - t_{sn-1} = \frac{Q * M}{S * r} * \tau * S = \frac{Q * M}{r} * \tau \quad (15)$$

Thus, the optimization of adjustment time is associated with the organization of periods of opportunities for students to take tests.

The motivation of the teacher and the student in organizing the organization of training tests with a sufficient number of passes is that students gain knowledge for the number of attempts (learning is excluded by the content, different types of tests, and their random mixing). Also, the teacher receives statistics on the verification of test questions. For this purpose, the teacher uses test questions on topics in colloquia or practical or laboratory classes.

If the teacher does not use the training tests, the adjustment phase increases significantly. Verification statistics will be obtained by the teacher only after the exam. This will mean that the correction of test questions will be a posteriori and its results can be used only for the following groups.

5. Conclusion

To quickly optimize the test parameters, it is necessary to provide the following conditions:

1. Form a test by topic and use it in the current learning process for a large number of students. This will allow to implement the test procedure many times and identify poor quality issues.
2. Such training tests have no restrictions on repeating the test.
3. The period of data preparation for their automated analysis is proportional to the total number of questions in the tests and inversely proportional to the number of students taking the tests, as well as the number of questions offered in the test.

Static tests that do not change and are not adjusted, especially on paper, have low suitability for optimizing their quality due to long periods of their application and the difficulty of accumulating sufficient data.

Exam tests should be formed based on a combination of training, verified at a sufficient level of tests.

The required quality of tests can be provided by special software modules that allow accumulating and process test data, to carry out a continuous process - "testing-processing-adjustment" for a short period of test quality assessment. Such software modules should have mathematical tools for analyzing the answers to the test questions, assessing the quality of the questions, the reliability of the results of the whole test, etc.

The testing module in the learning management system is one of the most complex. The testing process involves the active work of students and providing them with opportunities to feedback from the teacher on the results of testing. In addition, the accumulated statistical information allows you to analyze the results of assessment of students' knowledge and the level of quality of tests and their individual questions.

That is the general statistics on the use of test control during 01.09.2019 - 02.11.2020: tests total number - 3238, answers to the questions of electronic tests TestIQ: 302365. Given that only the first modules were conducted and there were no final tests, it can be concluded that the active use of training tests and the accumulation of analytics base for their adjustment. Training tests are the basis for the formation of examination tests.

The prospects for the development of the testing module are:

1. Modification of the test evaluation system.
2. Development and implementation of new procedural modules such as reminders on the timing of training tests, development of error test reports, etc.
3. Development of the concept of using elements of artificial intelligence in the testing modules.

6. Acknowledgments

We express our sincere gratitude for the activity and patience to the teachers and students of VNTU, who participated in the processes of developing and improving the quality of tests in the JetIQ system.

References

- [1] A. S. Serhii Maslovskiy, Adaptive test system of student knowledge based on neural networks, Proceedings of the 8th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS'2015), Warsaw, Poland 2 (2015) 940–944. doi:10.1109/IDAACS.2015.7341442.
- [2] N. Andriotis, Lms tools to help you assess your students' progress (2014).
- [3] A. Ostroukh, V. Blinova, T. Skvortsova, V. Nikonov, I. Ivanova, T. Morozova, Enhancement of testing process in learning management system moodle, Asian Journal of Applied Sciences 7 (2014) 568–580. doi:10.3923/ajaps.2014.568.580.
- [4] Collaborator, 2020. URL: <https://collaborator.biz/en/>.
- [5] Creating tests and managing test questions. learn how to create and manage a test in for your e-learning courses., 2020. URL: <https://www.docebo.com/knowledge-base/elearning-how-to-create-and-manage-test/>.
- [6] V. D. Rinaldi, N. A. Lorr, K. Williams, Evaluating a technology supported interactive response system during the laboratory section of a histology course, Anatomical Sciences Education 10(4) (2017) 328–338. doi:<https://doi.org/10.1002/ase.1667>.
- [7] J. West, W. Turner, Enhancing the assessment experience: Improving student perceptions, engagement and understanding using online video feedback., Innovations in Education and Teaching International 53(4) (2016) 400–410. doi:<https://doi.org/10.1080/14703297.2014.1003954>.
- [8] S. Gamage, J. Ayres, M. e. a. Behrend, Optimising moodle quizzes for online assessments., International Journal of STEM Education 27(6) (2019). doi:<https://doi.org/10.1186/s40594-019-0181-4>.
- [9] O. Kovalenko, Y. Palamarchuk, Algorithms of blended learning in it education, 2018 IEEE 13th International Scientific and Technical Conference on Computer Sciences and Information Technologies (CSIT) (2018) 382–386. doi:10.1109/STC-CSIT.2018.8526605.
- [10] Y. Palamarchuk, O. Kovalenko, R. Yatskovska, Variable assessment of students' knowledge using the test-iq system, Proceedings of the 9th scientific-practical conference. - Lviv: Publishing House of the Scientific Society. Shevchenko (2017) 188–193.
- [11] J. Rhode, S. Richter, P. Gowen, T. Miller, C. Wills, Understanding faculty use of the learning management system, Online Learning 21(3) (2017) 68–86. doi:10.24059/olj.v.vi.i.1217.