ISSN 2707-3092 (Online)
https://doi.org/10.32835/2707-3092.2020.20.160-166

Professional Pedagogics/1(20)'2020, 160-166 pp.

IVET Institute of vocational education and training of NAES of Ukraine

APPROBATION OF TEST TASKS AND ANALYSIS OF THEIR QUALITY FOR NON-STANDARDIZED TESTS

Andrii Kalenskyi[1], Oleksandr Yamkovyi[2]

1 Doctor of Pedagogical Sciences, Associate Professor, Head of laboratory of scientific and methodological support for specialist training in colleges and technical schools, Institute of Vocational Education of the National Academy of Pedagogical Sciences of Ukraine Kyiv, Ukraine
http://orcid.org /0000-0001-9034-5042, e-mail: kalenskyandrii@gmail.com

2 Candidate of Pedagogical Sciences, junior researcher of laboratory of scientific and methodological support of specialist training in colleges and technical schools, Institute of Vocational Education of the National Academy of Pedagogical Sciences of Ukraine Kyiv, Ukraine
https://orcid.org/0000-0001-6919-3515, e-mail: yamkovoysasha@ukr.net

Abstract.

Relevance. The issue of enhanced quality of professional junior bachelors training provides for a significant improvement in the control of student learning outcomes as an important means of educational process management. By means of pedagogical control, it is possible to discover the level of knowledge of graduates and obtain information about the state of students' knowledge in the educational process. The systematic feedback is provided through monitoring that enables to create an adaptive competence-oriented training program and in a timely manner correct activities of teachers and students in the educational process. Test control of knowledge as a method of measurement and assessment of knowledge, abilities and skills of a student most fully meets the requirements of legislation in the field of education to ensure relevance, transparency, reliability and objectivity of assessment of learning outcomes.

The purpose of this article is to highlight the methods of approbation of test tasks and analyze their quality for non-standardized tests.

Methods: of the theoretical analysis – to study existing approaches to the methodology of test development, test tasks and quality control of test tasks; expert evaluation – to determine the correct wording of test tasks and accuracy of the questions, time limit for fulfilling test tasks; approbation (pilot test) – to test the developed tools; Pearson correlation coefficient – to determine the correlation coefficient of task scores with the test total score.

Results: the concepts "test" and "test control" were defined, and the difference between standardized and non-standardized tests was determined, the specifics of expert evaluation of test effectiveness and its verification in the educational process in the context of trial testing was described, the algorithm for analysis of the test quality was revealed, the main indicators for selection of tasks for non-standardized tests and their calculation were singled out.

Conclusions. The article deals with approbation of non-standardized tests, their expert checks for correctness of the wording of the task, the accuracy of the wording of the questions, conditions for the implementation of trial testing; criteria for the selection of test tasks: the complexity potential; the variance of scores; the correlation coefficient of the task scores with the total test score. The further research will be focused on conducting a special experiment to create standardized tests, namely: check for reliability and validity based on a representative sample.

Key words: *test control, testing, test, test task, validity, reliability, complexity, variance of scores, correlation coefficient.*

Introduction. The quality of specialist training is compliance of learning outcomes in institutions of professional pre-higher education with legislation standards and requirements as well as the needs of employers and society. This quality is provided by external and internal education quality assurance systems of professional pre-higher education. According to the Law of Ukraine "On professional pre-higher education", one of the objectives of the internal education quality assurance system is to ensure the relevance, transparency, reliability and objectivity of assessment of learning outcomes.

The implementation of test control in the process of student knowledge assessment can meet the above requirements. However, the pedagogical practice shows that single-level tasks prevail in the structure of tests developed by teachers, the tests are developed without check for validity, reliability, and so on. Therefore, it can be noted that specific methodological recommendations are needed for construction of the pedagogical test and test tasks.

Sources. The issue of using tests in the educational process is not a new one. Many domestic and foreign scientists studied the issue of test control of student academic achievements. In particular, the theoretical and methodological aspects of test control were covered in the works of V. Avanesov, I. Bulakh, V. Bezpalko, V. Bocharnikova, N. Huliukina, V. Ilin, P. Luzan, E. Luzik, A. Maiorov, L. Morska, I. Romaniuk, L. Parashchenko, I. Pidlasyi, M. Chelyshkova, J. Fischer and others. The various forms of pedagogical tests and test tasks, recommendations for their use, principles and requirements for compilation of these tests are dealt with in these works. But despite there is a significant number of works dedicated to test control, the issue of test control methodology was not studied enough, in particular, there are no works that could justify detailed approaches to the construction of multi-level test tasks, and there is no methodology available for the average teacher to check test tasks and the test as a whole for validity.

Methods: of the theoretical analysis – to study existing approaches to the methodology of test development, test tasks and quality control of test tasks; expert evaluation – to determine the correct wording of test tasks and accuracy of the questions, time limit for fulfilling test tasks; approbation (pilot test) – to test the developed tools; Pearson correlation coefficient – to determine the correlation coefficient of task scores with the test total score.

The purpose of this article is to highlight the methods of approbation of test tasks and analyze their quality for non-standardized tests.

Results and discussions. At the current stage of reform of the professional pre-higher education system, the issue of test control of student knowledge and skills is becoming particularly important. According to A. Maiorov (2001, p. 14), a test is a tool consisting of a qualitatively verified system of test tasks, a standardized procedure for conducting and a pre-designed technology for results processing and analysis, intended to measure the qualities and features of a person, which can be changed in the process of systematic training.

L. Yaroshchuk (2010, p.51) defines the concept of a test as a combination of test tasks that mainly require an unambiguous answer, developed under certain rules and procedures, provides for a preliminary experimental test and corresponds to such effectiveness characteristics as validity and reliability, and test control of knowledge is a method of measurement and assessment of knowledge, abilities and skills of a student by means of specially prepared standardized tasks (Luzan and Kalenskyi, 2014, p. 62).

There are standardized and non-standardized test depending on the level of unification. Standardized tests are tests that were subject to a special experiment for standards development (criteria score systems), reliability and validity checks based on a representative sample, and have well-described instructions and methodology (Bulakh, 2005; Yaroshchuk, 2010, p.52).

However, in the activities of professional pre-higher institutions teachers, as a rule, use non-standardized tests, those compiled by themselves in certain academic subjects. This is due, firstly, to the fact that every academic subject taught in an institution of professional pre-higher education is constantly being improved, updated, and test tasks must be updated accordingly. Secondly, the use of non-standardized tests is predetermined by the following factors: much time spent by teachers on preparing tests and checking them for validity, time-consuming nature of the process, and the lack of a unified methodology for a practicing teacher to check test tasks and the test as a whole for validity. Therefore, in practice, ineffective test tasks are often used, which do not allow us to accurately determine the scope and quality of the educational material learned by the students. Thus, solution to the problem of test quality is based on the mathematical calculations of certain indicators, namely: complexity of the test task, variance of scores, correlation coefficient of the task scores with the total test score.

When a set of test tasks from a module or discipline section is developed, you must start its approbation and analysis. Approbation, or, as it is termed by

L. Burlachuk (2006), a pilot test, is conducted with a group of people who share characteristics with those for whom this test is intended. The purpose of approbation of test tasks is to:

1) determine the complexity of tasks and evaluate their suitability for students;

2) find tasks with significant shortcomings;

3) identify errors (spelling and punctuation mistakes, etc);

4) set the time limit to complete a task or a test as a whole;

5) analyze answers to open questions in order to specify the wording of correct answers;

6) identify shortcomings in the instructions.

Thus, in order to ensure the objectivity of data on the reliability of results obtained during the test control, it is necessary to check and confirm the validity of test tasks that will be included in the test of student learning outcomes assessment. Upon compiling a basic set of test tasks, it becomes necessary to check these tasks for relevance with the didactic goal to be reached. The effectiveness of the test is assessed through expert evaluation or comparison of testing results with assessments of other types of certification.

The first step towards test approbation is its expert evaluation. A group of teachers is formed based on the following criteria: academic degree and academic title, work experience in institutions of professional pre-higher education, taking into consideration their work experience as experts. The first expert review is made in order to assess the quality of a basic set of test tasks by expert teachers. The expert teachers assess the correctness of the wording of test tasks and the accuracy of questions, and set time for performing test tasks by themselves. The answers of expert teachers are compared with the reference ones prepared by the developers of test tasks. The remarks on the development of test tasks are discussed together with the developer and relevant adjustments and clarifications are made to their content.

For the final verification of the conditions for the development of test tasks, another expert review is made in the form of a pilot test involving a group of student experts. It is further determined whether students understand the task conditions and whether they understand the order of its fulfilment. In addition, the time they spend on completing a particular task is recorded, and their wishes are taken into account.

If the expert review showed that students are not able to fulfil the test within the time limit established in the structure of lessons for such a test check, the test needs to be revised, that is, the number is shortened, the complexity of tasks is adjusted etc. If the preliminary time limit is unknown, the experts determine the time necessary for fulfillment of tasks.

After completing the first version, the test is checked in the educational process as part of a trial testing (pilot test). The trial testing is the basis for collection of empirical material about the test quality. The statistical processing of this material enables developers to check the test for its validity and reliability. According to V. Avanesov (Avanesov, Khokhlova and Potap, 2005), in order to obtain statistically reliable results in trial testing, it is important to provide the following conditions:

1. Trial testing is to be held in several parallel groups. At the same time, it is recommended to conduct a trial testing twice in each group, provided that the test subjects receive two versions of the test with tasks that were not fulfilled before.

2. The number of subjects in groups should be at least 20 people.

3. All parallel groups are provided with the same conditions (time, place, duration of testing, etc.).

4. The time limit allotted for the trial testing is determined so that the most prepared students have time to answer all the questions of the test.

5. To get reliable results, it is necessary to minimize the possibility of sharing clues between the subjects of the test.

As soon as the trial testing is over, it is necessary to decide how to assess given answers to the questions. There are various models of assessment in the pedagogical literature. In our opinion, a dichotomous assessment model is the most convenient one: you will get one score for the correct answer (student ticked all correct solutions), but zero scores for the wrong answer (even if the student didn't tick at least one of all correct solutions, or, at least, one wrong solution in addition to correct ones). To check the test properties of tasks, the results of the trial testing shall be presented in a matrix format, which must be ordered. There are two orders in this matrix. The first of them relates to students who underwent testing, and the second refers to the lines with testing results, which are arranged in order of descending scores.

In the matrix of trial testing results, the lines with testing results are arranged in order of descending scores achieved by the students during testing: in the first line – surname of a student who achieved the highest score, and in the last – a student having the lowest score. We also use the above approach for fixing test tasks: from the easiest task to the most difficult one. The easiest test task is put on the first place, for which most correct answers were obtained, the task with less correct answers is put on the second, and so on, until the last one, for which only one correct answer was obtained. The scores achieved by all students in each of 10 tasks are indicated in the lower line of the table. The scores achieved by each subject of test is indicated in the last column of the table.

The matrix of testing results allows to make the first step in analyzing the quality of the test. For example, let's consider the matrix of testing results of knowledge of ten students *(table 1)*.

The main indicators for selection of test tasks for the test are the following: (Luzan, Kalenskyi and Kolisnyk, 2017): complexity potential, variance of

scores and correlation coefficient of the task scores with the total test score.

According to the experts, these indicators are the main criteria by which you can determine whether this task in the test form can be used in the test, or not. Therefore, it is necessary to make a number of statistical calculation*s (table 2)* in order to check the test properties of tasks in the test form and render some of them into test tasks.

The important step in test development is to check the *complexity* of the proposed tasks. Let's first consider how to determine the complexity potential of tasks.

The complexity potential in modern testing is calculated as the ratio of the number of wrong answers to tasks and the number of correct answers according to the formula:

$$P = q_i/p_i ,$$

where P is the complexity potential, pi is the percentage of correct answers of the subjects of test, and qi is the percentage of wrong answers.

As seen from the *table 2.8,* the first test task was answered correctly by all the respondents, and the tenth task, on the contrary, was not answered correctly by any student. After making calculations using the above formula, we get the following results: the complexity potential of task No. 1 equals to – 0 (0:10), and task No. 10 cannot be calculated (10: 0). Thus, tasks No. 1 and No. 10 should be removed from the list of test tasks that can be used when compiling tests.

The second important feature of test tasks quality is the *variance of scores*, which can serve as an indicator of the task differential ability, that is, the

Table 1
Example of a matrix of testing results of a group consisting of 10 students

S No.	Surname of the student	Task number										Σ
		1	2	3	4	5	6	7	8	9	10	
1	Haleta	1	1	1	1	1	1	1	1	1	0	9
2	Borysovych	1	1	1	1	1	1	1	1	0	0	8
3	Varlamenko	1	1	1	1	1	1	1	0	0	0	7
4	Avramenko	1	1	1	1	1	1	0	0	0	0	6
5	Dmytruk	1	1	1	1	1	0	0	0	0	0	5
6	Yanenko	1	0	0	0	0	1	1	1	1	0	5
7	Zakharenko	1	1	1	1	0	0	0	0	0	0	4
8	Kurinna	1	1	1	0	0	0	0	0	0	0	3
9	Smashko	1	1	0	0	0	0	0	0	0	0	2
10	Dontsiv	1	0	0	0	0	0	0	0	0	0	1
	Σ	10	8	7	6	5	5	4	3	2	0	-

Table 2
Analysis of test results

Indicators	Task number									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Number of correct answers	10	8	7	6	5	5	4	3	2	0
Number of wrong answers	0	2	3	4	5	5	6	7	8	10
Percentage of correct answers p_i	1	0.8	0.7	0.6	0.5	0.5	0.4	0.3	0.2	0
Percentage of wrong answers q_i	0	0.2	0.3	0.4	0.5	0.5	0.6	0.7	0.8	1
Complexity potential $P = q_i/p_i$	0.00	0.25	0.43	0.67	1.00	1.00	1.5	2.33	4.0	-
Variance of scores $S_j^2 = q_i p_i$	0	0.16	0.21	0.24	0.25	0.25	0.24	0.21	0.16	0
Correlation coefficient of task scores with total test score, r_{xy}	-	0.41	0.62	0.75	0.82	0.82	0.75	0.62	0.41	-

ability to divide a group of students into "strong" and "weak". It should be noted that the higher the variance of scores, the better the differential ability of the task.

For tasks which use only a dichotomous score (1 or 0), the variance is calculated by the formula:

$$S_j^2 = p_j q_j,$$

where p_j, q_j are the percentages of correct and wrong answers for each task. It is worth remembering that the higher the variance of scores, the better the differential ability of the task.

Determination of the correlation coefficient of task scores with total test score is the next statistical requirement for test tasks which is more accurate and technological measure of differential ability of the task.

To calculate the correlation coefficient, the Pearson product-moment coefficient is most often used:

$$r = \frac{n\Sigma xy - \Sigma x\Sigma y}{\sqrt{n\Sigma x^2 - (\Sigma x)^2}\ \sqrt{n\Sigma y^2 - (\Sigma y)^2}},$$

where r is the correlation coefficient; x is the result for each task; y is the score (result) for the entire test; n is the number of pairwise products; Σ – the summation sign.

For example, we calculate the correlation coefficient of scores on the seventh task (X_7). For this purpose, an auxiliary *table 3* is built up, where the corresponding data are used.

The first column shows values of the scores achieved by the students for the seventh task. The sum of these scores is 4 ($\Sigma X_7 = 4$). The second column shows the test scores achieved by each subject for all tasks, i.e. the total test score ($\Sigma Y_i = 50$). In the third column, the products of the scores of each subject for the seventh task (X_7) and sum of scores (Y) are indicated; $\Sigma X_7 \cdot Y = 29$. The fourth and fifth columns show the squared values X_7 and Y.

Respectively, $\Sigma X_7^2 = 4$ і $\Sigma Y^2 = 310$. To calculate the correlation coefficient, using the Pearson formula, we get:

$$r = \frac{n\Sigma xy - \Sigma x\Sigma y}{\sqrt{n\Sigma x^2 - (\Sigma x)^2}\ \sqrt{n\Sigma y^2 - (\Sigma y)^2}} = \frac{10*29 - 4*50}{\sqrt{10*4 - 4^2}\ \sqrt{10*310 - 50^2}} = 0{,}75$$

Tasks with a correlation coefficient more than 0.7 are considered to be key. Such tasks in the *Table 2* are those under number 4, 5, 6, 7. In particular, it is recommended to include only those tasks in the test that have a correlation coefficient at least 0.25 – 0.3.

The higher the rxy value, the more likely the task in the test form will be rendered into a test task, that is to be included in the test. The probability increases especially noticeably at $r_{xy} \geq 0.5$. If we calculate $(r_{xy})^2 x$ 100%, we get the value of the so-called coefficient of determination, expressed in a convenient for interpretation percentage of the correlation between the task and the total test score of the subjects. For example, the coefficient of determination in the seventh task is equal to $(0{,}8)^2 \cdot 100\% = 64\ \%$. It can be interpreted as follows: 64 % of the variation of test scores of test subjects in all tasks is related to the variation of scores on the seventh task. It demonstrates a high potential of the seventh task to make its contribution to the overall variance of the test.

The analysis of the test properties of the task also contributes to the calculation of a complete correlation matrix, which provides correlations of each task with all other tasks, as well as correlations with the sum of scores. This work is best done with the use of personal computers, by means of statistical packages.

Upon expert review and elimination of the identified shortcomings, a basic set of test tasks can be presented to students for use.

Table 3
Table for the correlation coefficient calculation

S No.	Surname of students	Score on task No.7, X_7	Total test score, Y_i	$X_7 Y_i$	X^2	Y^2
1	Haleta	1	9	9	1	81
2	Borysovych	1	8	8	1	64
3	Varlamenko	1	7	7	1	49
4	Avramenko	0	6	0	0	36
5	Dmytruk	0	5	0	0	25
6	Yanenko	1	5	5	1	25
7	Zakharenko	0	4	0	0	16
8	Kurinna	0	3	0	0	9
9	Smashko	0	2	0	0	4
10	Dontsiv	0	1	0	0	1
	Σ	4	50	29	4	310

Conclusions. The article deals with approbation of non-standardized tests, their expert checks for correctness of the wording of the task, the accuracy of the wording of the questions, conditions for the implementation of trial testing; criteria for the selection of test tasks: the complexity potential; the variance of scores; the correlation coefficient of the task scores with the total test score. The further research will be focused on conducting a special experiment to create standardized tests, namely: check for reliability and validity based on a representative sample.

Список посилань

Майоров, А. Н., 2001. *Теория и практика создания тестов для системы образования (Как выбирать, создавать и использовать тесты для целей образования*). Москва : Интелект-центр.

Ярощук, Л. Г., 2010. *Основи педагогічних вимірювань та моніторингу якості освіти : навч. посіб.* Київ : Видавничий Дім "Слово".

Лузан, П.Г. та Каленський А.А., 2014. *Методи і форми організації навчання фахівців аграрної галузі: навчальний посібник.* Київ : Видавництво "Логос".

Булах, І., 2005. *Основи педагогічного оцінювання. Частина І. Теорія: Навчально-методичні та інформаційно-довідкові матеріали для педагогічних працівників.* Київ : Майстер-клас.

Бурлачук, Л.Ф., 2006. *Психодиагностика : учеб. для вузов.* Санкт-Петербург: Питер.

Аванесов, В.С., Хохлова, Т.С., Потап, О.Е. и Плискановский, С.А., 2005. П*едагогические тесты. Методология разработки и применения: пособие для преподавателей.* Днепропетровск: Пороги.

Лузан, П.Г., Каленський, А. А. та Колісник, Н. В., 2017. *Технологія тестового контролю успішності навчання майбутніх молодших спеціалістів: методичні рекомендації.* Київ : ІПТО НАПН України.

Translation and transliteration

Majorov, A. N., 2001. *Teoriya i praktika sozdaniya testov sistemy dlya obrazovaniya (Kak vybirat, sozdavat i ispolzovat testy dlya celej obrazovaniya) [Theory and practice of creating tests for the education system (How to select, create and use tests for educational purposes)].* Moskva: Intelekt-centr [Moscow: Intellect Center], [in Russian].

Yaroshchuk, L.H., 2010. *Osnovy pedahohichnykh vymiriuvan ta monitorynhu yakosti osvity: navch. posib. [Fundamentals of pedagogical measurements and monitoring of education quality: training manual]*. Kyiv : Vydavnychyi Dim "Slovo" [Slovo Publishing House], [in Ukrainian].

Luzan, P.H. ta Kalenskyi, A.A., 2014. *Methody i formy organizatsii navchannia fakhivtsiv ahrarnoi haluzi: navchalnyi posibnyk [Methods and forms of training of specialists in the agricultural sector: training manual].* Kyiv: Vydavnytstvo "Lohos" [Kyiv: Lohos Publishing House], [in Ukrainian].

Bulakh, I., 2005. Osnovy pedahohichnoho otsiniuvannia. Chastyna I. *Teoriia: Navchalno-metodychni ta informatsiino-dovidkovi materialy dlia pedahohichnykh pratsivnykiv [Fundamentals of pedagogical assessment. Part I. Theory: Educational methodological and information reference materials for pedagogical staff].* Kyiv: Maister-klas [Kyiv: Maister-klas], [in Ukrainian].

Burlachuk, L.F., 2006. *Psihodiagnostika [Psychodiagnostics]* : ucheb. dlya vuzov. Sankt-Peterburg : Piter, [in Russian].

Avanesov, V.S., Hohlova, T.S., Potap, O.E. i Pliskanovskij, S.A., 200*5. Pedagogicheskie testy. Metodologiya razrabotki i primeneniya: posobie dlya prepodavatelej [Pedagogical tests. Development and Application Methodology: Manual for teachers. Dnepropetrovsk: Porogi [Dnipropetrovsk, Porohy]*, [in Russian]

Luzan, P.H., Kalenskyi, A. A. ta Kolisnyk, N. V., 2017. *Tekhnolohiia testovoho kontroliu uspishnosti navchannia maibutnikh molodshykh spetsialistiv: metodychni rekomendatsii [Technology of test control of success of training of future junior specialists: methodological recommendations].* Kyiv : IPTO NAPN Ukrainy, [in Ukrainian].

УДК 378.091.26:005.642.4

Апробація тестових завдань та аналіз їх якості для нестандартизованих тестів

Андрій Каленський[1], Олександр Ямковий[2]

1 доктор педагогічних наук, доцент, завідувач лабораторії науково-методичного супроводу підготовки фахівців у коледжах і технікумах, Інститут професійно-технічної освіти НАПН України, м. Київ, Україна

2 кандидат педагогічних наук, молодший науковий співробітник лабораторії науково-методичного супроводу підготовки фахівців у коледжах і технікумах, Інститут професійно-технічної освіти НАПН України, м. Київ, Україна,

Реферат.

Актуальність. Питання підвищення якості підготовки фахових молодших бакалаврів передбачає значне поліпшення контролю результатів навчання студентів як важливого засобу управління освітнім процесом. За допомогою педагогічного контролю розкривають рівень знань випускників та отримують інформацію про стан

знань студентів у освітньому процесі. Шляхом контролю здійснюється систематичний зворотний зв'язок, що дає змогу будувати адаптивну компетентнісно орієнтовану програму навчання та своєчасне коригування дій викладачів і студентів у освітньому процесі. Тестовий контроль знань як метод вимірювання й оцінювання знань, умінь та навичок студента найбільш повно відповідає вимогам законодавства у сфері освіти щодо забезпечення релевантності, прозорості, надійності та об'єктивності оцінювання результатів навчання.

Мета: висвітлення методики апробації тестових завдань та аналіз їх якості для нестандартизованих тестів.

Методи: теоретичного аналізу – для дослідження існуючих підходів до методики розроблення тестів, тестових завдань та перевірки якості тестових завдань; експертної оцінки – для визначення правильності формулювання тестових завдань, точності запитань і терміну виконання тестових завдань; апробація (пілотажне дослідження) – для перевірки розробленого інструментарію; коефіцієнта кореляції Пірсона – для визначення коефіцієнта кореляції балів завдання із сумарними балами тесту.

Результати: розкрито зміст понять "тест" та "тестовий контроль", визначено відмінність між стандартизованими та нестандартизованими тестами, охарактеризовано особливості здійснення експертної оцінки ефективності тесту та його перевірки в освітньому процесі в рамках пробного тестування, розкрито алгоритм аналізу якості тесту, визначено основні показники відбору завдань для нестандартизованих тестів та їх розрахунок.

Висновки: У статті розглянуто апробацію нестандартизованих тестових завдань, їх експертну перевірку на правильність формулювання завдань, точність формулювання запитань, умови здійснення пробного тестування; критерії відбору тестових завдань: потенціал складності; дисперсію балів; коефіцієнт кореляції балів завдання із сумарними балами тесту. Проте подальшими напрямами розвідок є робота з проведення спеціального експерименту для створення стандартизованих тестів, а саме: перевірка надійності та валідності на основі репрезентативної вибірки.

Ключові слова: *тестовий контроль, тестування, тест, тестове завдання, валідність, надійність, трудність, дисперсія балів, коефіцієнт кореляції.*

Received: 12 May 2020
Accept: 29 May 2020