УДК 378.146

DOI 10.31483/r-32751

Л.Ю. Лазарева, Т.С. Ларина, Т.В. Стебеняева

ОСНОВНЫЕ ТРЕБОВАНИЯ К МОДЕЛИРОВАНИЮ ПЕДАГОГИЧЕСКИХ ТЕСТОВ ДЛЯ КОНТРОЛЯ УРОВНЯ ЗНАНИЙ ОБУЧАЕМЫХ

***Аннотация**: в данной статье рассказывается о процессе информатизации сферы образования. Сегодня данный широкомасштабный процесс становится важнейшим фактором расширения практики внедрения инновационных технологий в образовательный процесс. Авторы статьи уточнили определение педагогического теста. В статье представлены состав требований к педагогическим тестам и их основные характеристики, выражающиеся через показатели надёжности, валидности и эффективности, а также раскрыта их сущность и приведены формулы для их расчета. Сделан вывод о том, что в условиях применения модульных образовательных технологий тестирование становится одной из эффективных форм контроля уровня знаний обучаемых.*

***Ключевые слова**: сфера образования, информатизация, обучение, модульные образовательные технологии, контроль уровня знаний, педагогические тесты, моделирование, экспертная оценка знаний.*

L.Yu. Lazareva, T.S. Larina, T.V. Stebenyaeva

MAIN REQUIREMENTS FOR THE MODELING OF PEDAGOGIC TESTS FOR THE PURPOSE OF' KNOWLEDGE LEVEL CONTROL OF THE STUDENTS

***Abstract**: the given article says about the process of IT development of education. Nowadays the said process becomes the most important factor of the extension of the practice of the implementation of innovative technologies in the educational process. The authors of the article gave the detailed definition of the pedagogic test. The article contains requirements for pedagogic tests and its main characteristics, expressed by such indices as reliability, validity and efficiency, as well as discloses*

its essence and provides formulas for its calculation. The opinion was made that in terms of the application of module educational technologies the testing becomes one of effective forms of the control of students' knowledge level.

***Keywords**: field of education, IT development, module educational technologies, knowledge level control, pedagogic tests, modeling, expert evaluation of knowledge.*

Nowadays the theory of modeling of pedagogic tests became, with good reasons, one of most important components of the educational process. In terms of the global IT development of the education field and of the extension of the practice of the application of module educational technologies (MET) the procedure of the evaluation of mastering by students of the didactic content of educational disciplines with the aid of pedagogic tests should be considered as the process of the objective measurement of the achieved knowledge level [3, p.87, 4].

In terms of the informational saturation of the educational process the test form of the control of knowledge considerably economizes the time and allows to overcome the subjectivism of evaluations of the level of students' knowledge, as well as enhances their motivation for the obtaining of the new knowledge. The application of the test control promotes the determination of the level of the knowledge mastering by students in the process of the whole period of studies of the one or another educational discipline [7, p. 154].

The basic purpose of this research is to determine basic requirements and relevant characteristics for the modeling of pedagogic tests and the control of students' knowledge level.

As MET is functioning in most higher education establishments, many qualified pedagogues interpret in the right the procedure of the control of students' knowledge with tests. With that, in order to avoid the ambiguous understanding of terms in this research, we will try to formulate the determination of the pedagogic test in this research more clearly.

First of all, it should be highlighted, that such English terminological notion as «test» is translated into Russian as «check», «verification». Besides that, in the modern literature there are multiple publications on this topic, representing all kinds of

real definitions of such term as «test» [8, 13, 14]. As a rule, it differs by the semantic content, but with that in practice it always includes a range of essential differences, attributed only to this term, namely:

1. the pedagogic test is considered by many experts as a specially generated test of specific tasks with increasing complexity;

2. the system of pedagogic tests is designed for the obtaining of quality evaluations of the level of students' knowledge by means of their performance of certain tasks [10, p. 69];

3. the system of pedagogic tests allows to measure effectively the level of mastering of the knowledge by students in the course of the educational process.

In addition, many pedagogues-scientists, while disclosing the essence of the studied term, extend its content, including the test structure and consequence as basic features (Efremova N.F., Strogonova E.I. and Mocropoulo A.A., Chourina K.V. and Zimina E.K [5, p. 27; 12; 15]). As more exact we see the interpretation of the essence of the term of «pedagogic test», given by V.S. Avanesov. Namely, he discloses the essence of this term as the system of mutually related tasks with the increasing complexity, which totality allows to evaluate the knowledge and skills, obtained by students in the process of studies of the didactic content of a certain educational discipline and its subject area [1, p. 31; 2].

Let's supplement the abovementioned determination with the following explanation. From the systemic approach positions the pedagogic test namely represents the specifically regulated system, consolidating different test tasks along functional lines. Different test tasks serve as basis for the formation of pedagogic tests' system in the whole totality of educational disciplines, studied in higher educational establishments.

In terms of the extension of the practice of application of MET in higher education establishments for the formation of quality tests it is necessary to observe certain requirements, which composition and content have been reflected in the specialized literature, dedicated to the computer testing problem. These papers pay a lot of atten-

tion to requirements for pedagogic tests. From the position of didactics following requirements are most significant ones:

1. the test content should correspond with testing purposes;

2. the test content should be comprehensive and systemic;

3. the content and forms of the representation of test tasks should be interrelated;

4. the test content should unambiguously determine the level of students' knowledge;

5. the test content should correspond to the modern level of development of such or another field of science [11, p. 20].

As a rule, in order to obtain results at use of pedagogic tests, it is necessary to carry out a range of measures, which are usually carried out with standard mathematic tests. The application of standard mathematic methods assumes the need of the correspondence of pedagogic tests' results with standard characteristics of the precision of measurement procedures. In conformity with the theory of pedagogic tests' modeling the quality level of the developed pedagogic should correspond to such precision characteristics of measurement procedures as *reliability, validity* and *efficiency* [2].

Characteristics of the test reliability are determined by stable indices, obtained in case of the need for reiterate measurements with the same test or its equivalent. It should be highlighted, that the increase of the content diversification of test tasks decreases the test reliability. The evidence from practice shows, that the use of the test at the verification of the mastering by students of the certain knowledge will also provide the more reliable result than the use of the test at the verification of such or another section (course) of the educational discipline. Such a circumstance takes place because the didactic content of even one section covers the considerably bigger scope of the material (conceptual provisions, regularities of the different origin, as well as direct facts) in comparison with the certain topic of the same section.

The reliability of pedagogic tests is considerably influenced by the complexity of its performance. Its value can be expressed by the proportion between correct and incorrect answers to test tasks, received from students. The total test reliability decreases sharply if: it comprises tasks, which are correctly answered by most students,

as well, otherwise, if in the test prevail tasks, incorrectly answered by the majority of students. Tests, comprising tasks, correctly answered by 50–85% of students, are most reliable and have the most practical value.

So, the essence of the characteristics of the test reliability reflects the precision of measurement and the stability of obtained results against occasional impacts of extraneous factors. That means that in case of minor changes of testing terms and of the condition of students during the testing final test results will undergo insignificant amendments.

The Cuder-Richardson formula is most widely spread for practical calculations of the coefficient of the test reliability (Rc):

$$R_C = \frac{k}{k-1}\left(1 - \frac{\sum_{j-1}^{k} p_j q_j}{\sigma_b^2}\right) \quad (1)$$

where R_C – is the reliability coefficient, k – the number of tasks in the test, p_j – share of correct replies to the jth task, q_j – the share of incorrect answers to the jth task, σ_b^2 – dispersion of individual grades of n students.

The following formula is usually used for the calculation of dispersion values:

$$\sigma_b^2 = \frac{\sum_{i=1}^{n} b_i - \frac{\left(\sum_{i=1}^{n} b_i\right)^2}{n}}{n-1} \quad (2)$$

where b_i – individual grades of the ith student.

As index of the high reliability of the test is considered the estimated value of the reliability coefficient R_C within 0,9 – 0,99. In such a case the evaluation of the test reliability can be considered as excellent. If the estimated value of the test reliability coefficient R_C is within 0,8 – 0,89, the student gets a good mark. If the estimated value of the reliability coefficient R_C is within 0,7 – 0,79, then its grade cannot be higher than satisfactory. And, finally, if the estimated value of the reliability coefficient R_C is

less than 0,69, i.e. its reliability is considered unsatisfactory, so, its following use in the pedagogic practice does not seem possible [9, p. 503].

The test reliability is usually enhanced by the increase of the quantity of included tasks. Anyway, actually it is rather problematic to develop the test with a large number of tasks, as there is always a whole range of different kinds of restrictions, namely: the time of the test performance, requirements for the observation of sanitary norms, time consumption, individual capabilities of students etc.

The content of characteristics of the test validity is close to the requirement for the fullness and versatility of the verification, the proportional representation of all elements of the studied didactic content of the topical area of such or such educational discipline. The «validity» term comes from the English word «valid», what is translated as actual, fit, applicable. The essence of these characteristics is that the test generation process should mandatory involve a qualified pedagogue, thoroughly familiar with all sections of the educational program, freely possesses skills of MET use in the educational process, well knowing how to set objectives and specified tasks for the mastering of knowledge by the contingent of students [6, p. 49]. Developed tests can become valid instruments for the evaluation of the level of knowledge of certain categories of students only if such terms have been met. If characteristics of the test validity are not indicated, it cannot be considered as a reliable instrument for the measurement and evaluation of the level of students' knowledge.

The clear definition of assignments within limits of the scope of knowledge within the subject area of the such or such educational discipline is the integrate term of the test validity. The test will not be active for students, to whom it is addressed, if included assignments are within limits of the scope of knowledge, mastered by students, as well as if it does not achieve these limits or exceed the scheduled knowledge level. In the pedagogic practice it is usual to highlight several kinds of the test validity, namely: conceptual, criteria, content, prognostic etc. The test validity is considered high enough, if the correlation coefficient (linear, Spearman's etc.) of correct answers and of the total number of tasks will be within 0,3 – 0,8.

Characteristics of the test validity serve as a certain index of its efficiency. Anyway, its value is almost never constant. It depends on students' contingent, terms of the test performance, character of the application of the mastered scope of knowledge in their future activity. So, there is good reason that for one situation a certain test can be found highly valid, and for another one – useless or even harmful. Characteristics of the test validity allow to interpret its results respectively to the testing purpose. In other words, the test validity demonstrates a certain measure of the test correspondence with targets, stated before the performance of the testing itself (use of a certain test).

Along with reliability and validity characteristics most important characteristics of pedagogic tests are its validity. The notion of «effectiveness» comes from the Latin «effectus», which can be translated as performance or action. So, the notion of «effectiveness» usually characterizes the result (effect), obtained after the performance of a certain action. As an effect mostly can be measured, its characteristics – efficiency, also refers to the number of measured notions. Usually it is determined through the relation between the very effect value and the value of expenses, time, resources etc.

Anyway, there is a material problem, arising at the evaluation of the efficiency of pedagogic tests, which essence is existing mutually excluding contradictions. In fact, on the one hand, it easily can be assumed that the test, which tasks' complexity fully corresponds to the students' knowledge level, should be considered as the most effective one. But, on the other hand, it also can be easily assumed, that the complexity of tasks of any test will ever correspond to the students' knowledge level, as each student has got personal abilities for studies and mastering of the new knowledge. This is why the evaluation of students' knowledge level requires a set of tests with tasks, having different complexity levels. Only this system of pedagogic tests can provide real evaluation of the students' knowledge level with enough high objectivity.

In order to evaluate the test efficiency its differentiating abilities are often used in practice. This index is considered higher, if in the process of the use of a certain test the are material discrepancies between sums of grades, received by students. Dispersion can be a possible index of differentiating abilities of the test. In such case

from two tests in the same educational discipline as most effective will be considered the one, having the bigger dispersion value. With the same number of tasks in both tests the proportion between the biggest dispersion value and its least value in per cents can serve as index of its comparative efficiency.

Provided the above mentioned it is necessary to disclose such notion as differentiated test efficiency. It appears to us that no test can be considered efficient without the whole period of students' preparation. The efficiency of each certain test is determined by the totality of the didactic content of the knowledge, for which the set of included tasks fully corresponds to the established complexity level. Consequently, each certain test can be effective for a one level of knowledge and not be such for another level of knowledge. So, in order to obtain objective evaluations of the students' knowledge level it is necessary to generate the system of effective tests, each of which will be the optimal instrument for the measurement of the students' knowledge level at certain stages of their studies.

In the course of performed research following opinions have been made:

1. The evaluation of students' knowledge level with the test control has got the important educational & developing value, as it promotes the comprehensive studies of the didactic content of the subject area of an educational discipline, the extension and intensification of volumes of the students' knowledge, as well as upgrading and development of their cognitive interests.

2. The test control, as one of universal instruments of pedagogic measurements of the students' knowledge, lays the groundwork for the generation of the effective system for the education quality measurement.

3. The developed technologic education basis makes the testing not only as a mean for the control and evaluation of students' knowledge level, but also stimulates them to the independent educational activity.

4. Nowadays effective test systems in different subject areas of studied educational disciplines become the integrate part of the educational process in any higher education establishment.

The research has been performed with the partial financial support of the Russian Fund of Fundamental Research, project №17–06–00010-ОГН-а «Development of the Theory and Instruments for the Testing of Students' Knowledge Level in Terms of the Comprehensive Application of Electronic Educational Resources ».

Список литературы

1. Аванесов В.С. История педагогической теории измерений // Образовательные технологии. – 2012. – №4. – С. 28–38.

2. Аванесов В.С. Эффективность педагогических тестов и тестовых заданий [Электронный ресурс]. – Режим доступа: http://www.testolog.narod.ru (дата обращения: 20.04.2019).

3. Белобородов В.Н. Применение современной теории тестирования IRT в системе контроля измерительных свойств диагностических материалов / В.Н. Белобородов, А.О. Татур // Педагогические измерения. – 2016. – №2. – С. 85–97.

4. Говорова И.Е. Применение тестовых технологий в рамках реализации требований ФГОС [Электронный ресурс]. – Режим доступа: http://www.informio.ru/publications/id1967/Primenenie-testovyh-tehnologii-v-ramkah-realizacii-trebovanii-FGOS (дата обращения: 20.04.2019).

5. Ефремова Н.Ф. Критериальные требования к фондам оценочных средств // Педагогические измерения. – 2016. – №1. – С. 25–31.

6. Ларин С.Н. Обоснование использования новых методов контроля уровня знаний обучаемых в составе модульных технологий обучения / С.Н. Ларин, Л.Ю. Лазарева, Г.С. Худолей // Международный научно-исследовательский журнал. – 2018. – №1(67), ч. 4. – С. 48–51.

7. Ларин С.Н. Организация контроля уровня знаний обучаемых на основе технологий тестирования / С.Н. Ларин, Т.С. Ларина, Е.Ф. Баранов // Успехи современной науки и образования. – 2017. – №6, Т. 2. – С. 152–156.

8. Ларин С.Н. Современные подходы к моделированию тестов: система требований, преимущества и недостатки, основные этапы разработки / С.Н. Ларин, У.Х. Малков // Мир науки. – 2016. – Т. 4, №1 [Электронный ре-

сурс]. – Режим доступа: http://mir-nauki.com/PDF/04PDMN116.pdf (дата обращения: 20.04.2019).

9. Печников А.Н. Решение задач текущего педагогического контроля на основе анализа результатов критериально-ориентированного тестирования / А.Н. Печников, Д.А. Печников // Образовательные технологии и общество. – 2015. – Т. 18, №2. – С. 489–513.

10. Соколова И.И. Дидактический потенциал и риски тестирования // Человек и образование. – 2017. – №1 (50). – С. 68–72.

11. Стебеняева Т.В. Инновационные технологии и формы контроля уровня знаний обучаемых / Т.В. Стебеняева, Е.Ф. Баранов, Г.С. Худолей // Научное обозрение: гуманитарные исследования. – 2015. – №1. – С. 17–26.

12. Строгонова Е.И. Адаптивная модель контроля и оценки знаний, обучающихся в условиях компетентностного подхода / Е.И. Строгонова, А.А. Мокропуло // Конкурентоспособность в глобальном мире: экономика, наука. – 2016. – №3.

13. Устинова И.Г. Применение тестирующих программ в современном образовательном пространстве / И.Г. Устинова, Е.Г. Лазарева [Электронный ресурс]. – Режим доступа: http://portal.tpu.ru/f_dite/conf/2014/1/c1_Ustinova.pdf (дата обращения: 20.04. 2019).

14. Феськов Н.С., Якобчук А.П. Педагогический тест как инструмент объективного измерения уровня подготовки / Н.С. Феськов, А.П. Якобчук [Электронный ресурс]. – Режим доступа: http://www.giac.unibel.by/modules.php (дата обращения: 20.04.2019).

15. Чурина К.В. Тестирование как форма контроля результатов обучения / К.В. Чурина, Е.К. Зимина // Молодой ученый. – 2015. – №9. – С. 1214–1217 [Электронный ресурс]. – Режим доступа: https://moluch.ru/-archive/89/18283/ (дата обращения: 20.04.2019).

References

1. Avanesov, V. S. (2012). Istoriia pedagogicheskoi teorii izmerenii. *Obrazovatel'nye tekhnologii, 4*, 28-38.

2. Avanesov, V. S. Effektivnost' pedagogicheskikh testov i testovykh zadanii. Retrieved from http://www.testolog.narod.ru

3. Beloborodov, V. N., & Tatur, A. O. (2016). Primenenie sovremennoi teorii testirovaniia IRT v sisteme kontrolia izmeritel'nykh svoistv diagnosticheskikh materialov. *Pedagogicheskie izmereniia, 2*, 85-97.

4. Govorova, I. E. Primenenie testovykh tekhnologii v ramkakh realizatsii trebovanii FGOS. Retrieved from http://www.informio.ru/publications/id1967/Primenenie-testovyh-tehnologii-v-ram

5. Efremova, N. F. (2016). Kriterial'nye trebovaniia k fondam otsenochnykh sredstv. *Pedagogicheskie izmereniia, 1*, 25-31.

6. Larin, S. N., Lazareva, L. Iu., & Khudolei, G. S. (2018). Obosnovanie ispol'zovaniia novykh metodov kontrolia urovnia znanii obuchaemykh v sostave modul'nykh tekhnologii obucheniia. *Mezhdunarodnyi nauchno-issledovatel'skii zhurnal, 1(67), ch. 4*, 48-51.

7. Larin, S. N., Larina, T. S., & Baranov, E. F. (2017). Organizatsiia kontrolia urovnia znanii obuchaemykh na osnove tekhnologii testirovaniia. *Uspekhi sovremennoi nauki i obrazovaniia, 6, t. 2*, 152-156.

8. Larin, S. N., & Malkov, U. Kh. (2016). Sovremennye podkhody k modelirovaniiu testov: sistema trebovanii, preimushchestva i nedostatki, osnovnye etapy razrabotki. *Mir nauki, T. 4, 1*. Retrieved from http://mir-nauki.com/PDF/04PDMN-116.pdf

9. Pechnikov, A. N., & Pechnikov, D. A. (2015). Reshenie zadach tekushchego pedagogicheskogo kontrolia na osnove analiza rezul'tatov kriterial'no-orientirovannogo testirovaniia. *Obrazovatel'nye tekhnologii i obshchestvo, T. 18, 2*, 489-513.

10. Sokolova, I. I. (2017). Didakticheskii potentsial i riski testirovaniia. *Chelovek i obrazovanie, 1 (50)*, 68-72.

11. Stebeniaeva, T. V., Baranov, E. F., & Khudolei, G. S. (2015). Innovatsionnye tekhnologii i formy kontrolia urovnia znanii obuchaemykh. *Nauchnoe obozrenie, 1*, 17-26.

12. Strogonova, E. I., & Mokropulo, A. A. (2016). Adaptivnaia model' kontrolia i otsenki znanii, obuchaiushchikhsia v usloviiakh kompetentnostnogo podkhoda. *Konkurentosposobnost' v global'nom mire, 3*.

13. Ustinova, I. G., & Lazareva, E. G. Primenenie testiruiushchikh programm v sovremennom obrazovatel'nom prostranstve. Retrieved from http://portal.tpu.ru/f_dite/conf/2014/1/c1_Ustinova.pdf

14. Fes'kov, N. S., & Iakobchuk, A. P. Pedagogicheskii test kak instrument obektivnogo izmereniia urovnia podgotovki. Retrieved from http://www.giac.unibel.-by/modules.php

15. Churina, K. V., & Zimina, E. K. (2015). Testirovanie kak forma kontrolia rezul'tatov obucheniia. *Molodoi uchenyi, 9*, 1214-1217. Retrieved from https://mol-uch.ru/archive/89/18283/

Лазарева Лариса Юрьевна *– канд. техн. наук, ведущий специалист АНО ДПО «Институт международных стандартов учета и управления», Россия, Москва».*

Lazareva Larisa Yurievna *– candidate of technical sciences, lead expert at the ANO CPE «Institute of International Standards of Accounting and Management», Russia, Moscow.*

Ларина Татьяна Сергеевна *– научный сотрудник АНО ДПО «Институт международных стандартов учета и управления», Россия, Москва.*

Larina Tatiana Sergeevna *– researcher at the ANO CPE «Institute of International Standards of Accounting and Management», Russia, Moscow.*

Стебеняева Татьяна Викторовна *– канд. экон. наук, ведущий специалист АНО ДПО «Институт международных стандартов учета и управления», Россия, Москва.*

Stebeniaeva Tatiana Victorovna – *candidate of economic sciences, lead expert at the ANO CPE «Institute of International Standards of Accounting and Management», Russia, Moscow.*