

# Methodological Principles and Methods of Design and Evaluation of Education Tests in History Education

Albina R. Fayzullina<sup>1</sup> & Firuza A. Saglam<sup>2</sup>

<sup>1</sup> Kazan (Volga Region) Federal University, Kazan, Russia

<sup>2</sup> Institute of Economy, Management and Law, Kazan, Russia

Correspondence: Albina R. Fayzullina, Kazan (Volga Region) Federal University, 420008, Kazan, Kremlyovskaya Street 18, Russia. E-mail: fazalbina@rambler.ru

Received: June 15, 2015 Accepted: June 24, 2015 Online Published: June 30, 2015

doi:10.5539/jsd.v8n7p35

URL: <http://dx.doi.org/10.5539/jsd.v8n7p35>

## Abstract

The article is relevant due to the insufficient development of the problem of objective and independent assessment of the students' academic achievements level in the theory of education and the complexity of optimum combination in history education. The aim of the article is to analyze the principles of design of tests and education tests which monitor historical knowledge of all types and levels. The main method of the study of this problem was criterion oriented approach that allows to realize the possibility to design and use tests of different types for pedagogical monitoring on History. The article deals with the basic concepts and terminology of the theory of development and use of educational tests as well as the stages of test design. The article also deals with the types of pedagogical control in which the tests such as current, thematic, mid-term and final tests are used. Recommendations for selection of items of the content of the subject for tests in accordance with the aim of testing and students' number are given. The materials of the article might be useful for the teachers of History to design educational tests for academic aims and analyze made-up evaluative (monitoring and measuring) materials on History.

**Keywords:** pedagogical testing, education test, test specification, test item, test item form

## 1. Introduction

The evaluation of students' academic achievements is important and mandatory part of the process of practical activity of any teacher.

Academic achievements in a broad sense are considered as the set of subject knowledge, ideas, skills and competences formed in the process of systematic education in the educational establishment (Khayrutdinov, 2014) or during self study according to approved program. The basic functions of the pedagogical control are diagnostic, monitoring, educating, developing and prognostic (Lerner, 1982; Levina, et al, 2015; Khanmurzina et al, 2015; Ivanenko, et al, 2015).

Diagnostic, monitoring and prognostic functions dominate when using test forms of pedagogical control. Namely these functions are realized during the control at the different stages of continuous educational process.

Entrance test is used before the start of new course, new section or topic of the subject. According to it the basic level of knowledge necessary to understand and master the content of academic subject is determined (Efremova, 2003).

The current test is held during the education process to update and consolidate the students' knowledge. It allows the teacher to discover the problems of each student's knowledge, to adjust and restructure the education process in time, and this test provides feedback in the "teacher-student" system (Mayorov, 2000).

The aim of the thematic test is to evaluate the level of students' academic achievements in certain spheres of knowledge which corresponds with the finished section or topic of the course.

Analyzing data received during the process of thematic monitoring of the training group (class), the teacher sees the results of education process and decides if it is possible to move to the next section of the program or it is necessary to have additional lessons on the topic. Thematic test allows to get the information about the dynamics of mastering the materials by the training group in general and by each student in particular (Andersen, 2011).

In the system of continuous monitoring of the quality of educational process it is very important.

Mid-term test helps to discover the results of specific temporary stage of education (term or semester, year or period of education). Final test is held on the stage of transfer from one level of education to another and must correspond to the compulsory requirements to the training quality on the subject which are determined according to the educational standards (Pereverzev, 2003).

So, in the secondary school three levels of education are determined on finishing of which the final test is held. These are junior school, middle school and senior secondary school.

As any type of pedagogical monitoring, tests also have educational, motivating and developing functions. Thematic and mid-term tests show the student his level achieved during the education and force him to make adequate actions to adjust his knowledge and skills in the process of education. The results of the final test allow to make a decision about further education, for instance, about applying to higher educational institutions or colleges. These functions are actively used in the students' self-study with the tests. In many higher educational institutions, for example, student can realize self-control on any section of the program with the help of computer tests (Kadnevsky, 2004).

## **2. Methodological Framework**

### *2.1 Research Objectives*

The following objectives were to be worked out in the process of the research.

- 1) Characterize methodological principles and the techniques of pedagogical tests design in history education;
- 2) Justify the variants of evaluation of pedagogical tests.

### *2.2 Theoretical and Methodological Basis of the Research*

Theoretical and methodological basis of the research were conceptual thesis of pedagogics related to the determination of the notion of pedagogical testing, education test, test specification, test item and test form. To analyze the highlighted problem such approaches as institutional and structural and functional were used, as well as general scientific and special methods which are the tools of pedagogical science, in particular these are comparative and contrastive. The selection of the elements of the content was realized on the bases of State educational standards, programs of basic and enhanced levels of learning History at school and university, programs for the applicants, text books and schoolbooks which are recommended by the Ministry of education and science of Russian Federation.

## **3. Results**

The analysis produced by M.B.Chelyshkova (2001) shows that education tests might be used in any types of pedagogical monitoring. But apart from the traditional means of control, tests allow to discover not only the level of training, but the structure of the students' knowledge, namely the degree of digression from that structure which was planned by the teacher.

Pedagogical testing is a combination of organizational and methodological events united by the general aim with education tests and intended to train and hold official procedures of test delivery, processing, interpretation and presenting the results on accomplishment.

According to N.F. Efremova (2010) education test is a system of specially selected evaluative items of specific (testing) form that allows to evaluate the academic achievements in one or several areas of knowledge.

There are so many classifications of education tests (Michailychev, 2002). The bases for their classification might for instance be:

- the way of test delivery: blanks, test copybooks, computer;
- number of subjects, knowledge of which are tested: homogeneous test allows to evaluate the level of training on one specific subject, heterogeneous test allows to evaluate the level of training on several allied subjects;
- number of different test forms of test items: monomorphic test contains test units of one form, polymorphic test contains the test items of various forms and types;
- the difficulty of items included into the test: the capacity of test consists of test items of increasing difficulty, pace of test contains the items of the same difficulty;
- the structure of test: with the location of test items according to the logics of the subject or according to the increasing difficulty of test items;

- the type of pedagogical monitoring: entrance test, current test, thematic test, mid-term test, final test;
- the aim of the test result usage: assessment test, enrollment test, current test;
- the way of test delivery: blank test (student is given printed version), subject test (student needs to manipulate with material objects), and computer tests;
- methodology of interpretation: criterion-oriented and standard-oriented;
- the range of usage: for the teachers in the current education process, for the administrative body of the educational institution to assess teachers, centralized usage as large samples to select and form the groups;
- the professionalism of the developer: tests are standardized and informal.

V.S. Avanesov (1998) points out several stages in developing education tests.

1. The determination of the testing aims (form of control, number of students, resources).
2. The analysis of the content of the subject and selection of studying materials, the knowledge of which is monitored by the test.
3. The development of the test specification.
4. The designing of test items in accordance with the specification, the configuration of tentative test.
5. The approbation and review of tentative test.
6. Statistical processing of the approbation results.
7. The adjustment of test items, the exclusion of unsatisfactory items and selecting the satisfactory items for the final version of the test.

Is it obligatory to fulfil all the mentioned stages? The answer to this depends on the aim of the test usage. For instance, if the test is developed by the teacher to evaluate the level of students' training on the specific part of the subject (for example, for the current monitoring) and is used only by the developer in some specific education process, then some stages might be omitted (for example, the development of the test specification, approbation, statistical processing). Because the teacher in this case needs only to discover the percentage of the students of the class who obtained certain knowledge and skills, which were planned by the teacher and are necessary for the students to understand further educational material. It is obvious that even in this case it is required to make a plan of the test at first, the table in which each test item is related to a certain element of the subject content and certain type of knowledge or skills that are to be monitored with the help of this test item.

If the test is intended for the multiple uses among large number of students, for example, for the final test of the school pupils of the 9<sup>th</sup> grade with the aim of monitoring the quality of history teaching, then the fulfilment of all the above mentioned stages of the test developing are obligatory. The test for mass final testing, like centralized testing, the State unified exam are developed more thoroughly (Michaleva & Khlebnikov, 2000).

The most significant stage that influences all the other stages is the first one, which is the stage of goal-setting. It is very important to determine for what type of pedagogical monitoring this test will be used and how the results of the test will be interpreted at the very beginning of the test developing.

The most important performance goal of the test is the methodology of interpretation: the criterion-oriented and standard-oriented tests are differentiated.

The criterion-oriented tests are intended to evaluate the level of training of each student according to the requirements of the education program or its part. Individual result in such kind of testing is compared with the previously planned result (criterion), and not with the achievements of the other students. So the aim of the criterion-oriented test is the assessment of the student in accordance with the level of mastering the certain area of the content.

The area of the content is displayed in the criterion-oriented test as the system of items on the maximum number of the elements of the program content or its part. The level of the specification of the content is thorough.

The most important thing in developing the criterion-oriented test is the clear correspondence of the number, content and the difficulty of items with the monitoring requirements of the program. That's why the items with which none of the students can cope or all can cope are not excluded from the test, if they are important elements of the content according to the program. Test score for such tests are determined by the percentage of correctly accomplished items.

All the assessment tests, entrance tests, thematic and mid-term tests are considered to be the criterion-oriented tests (Pereverzev, 2003).

Standard-oriented test is intended for the certain level of training of the students to compare their results with each other and/or with testing standards. Hence we have the name of the test. Testing standard is average statistic value of the test score, determined for this test. It is previously determined according to the results of approbation of test at the representative sampling of the students or after the mass test is held, if the representative sampling is impossible to be set beforehand. For instance, if the testing is mandatory for all the pupils or voluntary, then it is impossible to calculate how many students with different level of training on the subject will come to the testing.

The most important thing in developing standard-oriented test is the inclusion of such items, the accomplishment of which allows in a maximum way to differentiate the students according to the level of training. The items of medium difficulty have the highest differentiating capability and that's why they dominate in such kind of tests. To determine whether it is so or not and if the test items correspond with the aims of standard-oriented test (differentiation of the students by the level of training), the approbation of the test is held at representative sampling.

In this case the items with which none of the students can cope or all can cope are not included into such test. The level of specification of the content area of the test is not so important. If the test items reflect the most significant elements of the content, this would be enough. That's why standard-oriented tests as a rule contain fewer items than the criterion-oriented tests.

Standard-oriented tests are those intended for the competitive selection to get an education in the universities in particular (for example, enrollment tests of the centralized testing).

After setting the goals of the testing, the analysis of the subject content and regulatory requirements to the level of students' training is made, knowledge and skills which will be tested and the level of mastering of them are determined.

The result of the first two stages of the pedagogical testing development is the specification of the test, that is a document in which there is information about the aim of testing, the content of the subject and the types of knowledge and skills, monitored by test items, as well as the main requirements for the regulations to hold a test, processing the test results and their interpretation. Specification of test usually contains the list of normative documents which determine the selection of test item content in accordance with the aim of testing, list of references necessary to train for the testing (Chelyshkova, 2011).

The integral part of specification is the plan of the test, which is a table where each test item corresponds with the certain element of the subject content and monitored knowledge or skill, level of difficulty, form and type of the test item. There are examples of instructions to the test items of different form in the specification of the test. So the specification is the full and comprehensive description of the test.

Education test consists of test items. The quality of test items determines the quality of education test.

Test item is a minimum meaningfully completed part of education test in the form of test.

Test form consists of instruction to the test items, text of task and/or non-verbal materials and the system of evaluation of the test accomplishment. The non-verbal materials are pictures, figures, maps, tables, which might be included into the test.

The instruction to the test contains the guidance on the ways of registration of the test items accomplishments, in particular the ways of writing correct answers (what, how and where to tick, write), writing or design of the mathematical problem solution and so on. The instruction might be the same for several test items if they are of the same type.

Modern scientists and testologists M.B. Chelyshkova (1995), A.N. Mayorov (2000), V.S. Avanesov (2005), V.Yu. Pereverzev (2005) summarized the principles of selection of the test items content:

- of consistence, comprehensiveness and equilibrium of the test content which reflect the requirements to the consistency, generalization of the content of knowledge, link of the test items with the logical structure of educational materials. If to follow the principles of consistency the test might be used to evaluate not only the content of knowledge, but the quality and the structure of students' knowledge;
- of congruence which is the correspondence of the education test with the content of academic subject, and it is necessary for the test items to embrace the most important aspects of the content area in the right proportion;

- of significance which is the necessity of inclusion into the test only the most important, basic knowledge. This knowledge in which the essence, the content, the laws and regularities of the reviewed phenomena are reflected;
- of scientific certainty, which means that the tests include only the true and scientifically justified subject content. All disputable viewpoints that are admissible in science are excluded from the test items;
- of correspondence of the test content with the level of modern state of scientific knowledge. The difficulty is in some stoppage which is necessary to adjust scientific achievements and describe them in the form that might be understood by the students;
- representativeness which means more accurate reflection of necessary knowledge in the test items including necessary links within the subject. As many elements of knowledge are linked between each other, hierarchically are subdominant, included one into the other, it is necessary to choose monitored elements of the content, to include “the structural items” into the test monitoring the comprehensive knowledge;
- of flexibility of the content which means that the test content might be permanent and must depend on the development of the science, the change of the academic programs and the release of new books and so on;
- of interconnection of the content and the form which means to control only those elements which might be expressed in test form.

The text of the items and non-verbal materials to them are the meaningful bases of the test items. The content of the test items are determined by the structure of the subject content, aim of the testing and the type of testing and must correspond with the test plan. It is important that each test item corresponds with the content of the certain type of knowledge (skills). It is specifically important when the test is developed to diagnose the structure of the students' knowledge and the quality of the education process technology. The results of such testing would help to understand what namely and in what level the students mastered. While developing the current and mid-term tests this requirement is fulfilled easily. The main function of the final test is not the diagnosis of the structure of the students' knowledge, but control and forecast of the readiness to further education. In the final test, especially with the aim of competitive selection, one test item monitors the integrative (all-inclusive) knowledge and skills of several (usually not more than two) elements of the content.

#### 4. Discussions

The specific features of academic subject “History” according to a well-known professional on the history teaching methods I.Ya. Lerner (1982) “considers the presence of knowledge, skills, the experience of creative activity and the experience of the emotional and sensitive attitude to the historical and social phenomena. These elements with all their relative self-sufficiency are closely interrelated. There is no knowledge without skills and so the experience of realizing the ways of activity is acquired on the bases of knowledge. The creative activity is always meaningful, and it is realized with the help of knowledge and skills <...>. In other words, every other type of content relies on the previous as on its bases” (Lerner, 1982).

On the bases of this thesis, most test items on history have complex and integrative character. It is important to manage to single out in every test item the most significant element of graduate training and to compose the text of the test item in such a way that it would check and in every specific case namely this, primary element of the content and certain type of activity.

Test items are elementary units of the test, its “bricks”. But the test is not a set of test items, but the system of test items. It means that every test item should have certain features, which make them the element of system, carrying out the fixed aim of pedagogical monitoring. Such features of test items are firstly, quality characteristics which are required by the test specification: controlling element of the content, the form of testing items, the type of monitored knowledge and skills, required level of mastering, the way of evaluating the item accomplishment. These characteristics and their correspondence with specification are evaluated by expertise. Secondly, quantitative or statistical characteristics of the test items which embrace the difficulty and differentiating power of the test item. These characteristics of each test item are discovered during the approbation (empirical verification) of test in general as the system of test items, and not as every specific test item.

According to the results of approbation and inspection of the test some parts of test items might be adjusted or totally substituted, if it is clarified that some test item is incorrect in formulating or too easy or visa versa too difficult. For standard-oriented tests intended for mass testing it is sometimes required to repeat the procedures: approbation, inspection, adjustment of test items. As a result of this test becomes more appropriate for the aim, for which it is developed, and test items became systematically important elements. Only in this case the test

items become fully test items of the certain test. In this respect all the test items in the test which didn't get approval might be considered as pre-test items.

The basic difference of the test from the traditional form of control is the possibility to assess the students' level of training objectively, quantitatively evaluate his knowledge and skills according to the certain amount of the elements of the subject content. In this case the score for the test accomplishment depends on the score of each test item accomplishment in the test. There are two types of test items according to the ways of evaluating: dichotomous and polytomous (Michailychev, 2003).

The accomplishment of the dichotomous test item is evaluated only alternatively (1 point if it is correct, and 0 point if it is incorrect).

The accomplishment of the polytomous test item allows several variants of answer, each of which evaluated differently. For example, one of the category of answers might unite the results of one or several stages ("steps") of problem solution. For History, or other humanitarian subjects, one category of answer more often unites different correct but not full (partially correct) answers. So the category of answers is a combination of variants of answer for the test item. Depending on the quality of an answer, the score corresponding with the certain category of answer for the polytomous test item will be different. Category of not full or partially correct answer might be given 0 point or 1 point, and category of full answer will be given 2 points.

If for example in test items on History it is required to choose from the given list, which contains 5 points, the characteristics of any historical event (period or personality) and three of them are related to the mentioned event, the system of evaluation of this test item might be the following:

- for totally correct indicated characteristics 3 points;
- for two correct indicated features 2 points;
- for one correct indicated feature 1 point,
- not any correct feature 0 point.

In this case it is said that polytomous test has four categories of answer. For the same test item there might be used more strict system of evaluation:

- for 3 correct features 2 points;
- for 2 features 1 point;
- all other variants of answer are evaluated as 0 point.

In such system of evaluation polytomous test item has three categories of answer.

Dichotomous test item can be considered as polytomous test item with two categories of answer.

Dichotomous system of evaluation of test item accomplishment is reasonable to use if the test in general has a great number of test items and to accomplish the test a lot of time is given. In this case polytomous evaluation will make it harder to check students' works. Dichotomous evaluation is used in mid-term and final monitoring of knowledge as well as in cases when it is necessary to check a lot of answer sheets manually. For instance, thirty works.

Polytomous system of evaluation should be involved if the test items control understanding of large content elements, which might be structured or divided into small parts (features, events, facts, characteristics and so on). In this case one such test item gives a lot of information about student's knowledge, as it embraces large part of education material. Polytomous test items can be used in all types of pedagogical monitoring. The current and thematic tests, when short period of time is given to fulfil the test, can consist of only several polytomous test items.

In the final tests when the time to fulfil the test is more (two or three hours) and the total number of test items in the test might be about fifty, the number of test items with the polytomous system of evaluation might be five or ten. As the experience of such test items usage in the final centralized test showed, the accuracy of the level of training assessment increase, but the technology of result processing becomes more complicated, so the price to develop the test and result processing increase.

So the system of test item evaluation must not be a goal in itself. The choice of the system of evaluation, first of all, is defined by the aim of testing and the type of pedagogical monitoring. It also depends on the type and form of test items.

The determination of the test item form which was mentioned above is a bit in simplified form, as it contains general features of test items. However, any test item in the testing form must correspond with the certain set of specific requirements, the fulfilment of which provides universal understanding of the problem (test item) and excludes the possibility of wrong answer emergence on formal features. The system of evaluation must provide single-value evaluation of test item accomplishment.

For any content element determined to be included into test, it is necessary to set the level of its being mastered and the type of knowledge and skills which student should obtain during the education. This is rather important moment in planning the test from the point of view of activity and competency-based approaches. Otherwise the education will be mainly oriented on the formation of knowledge of the first level of mastering (reproduction of the content by memory) and only factual knowledge will be monitored.

While developing the test it is essential to remember that the result of the history education must be mastering specific types of activities for this subject and general educational activities related to different levels of mastering academic material. Test items must control all these types of activities in maximum amount. That's why the developer needs in their detailed list (Emelyanov & Yatsenko, 2002).

The clarification of the type of knowledge and skills which are formed while learning this or that section (subsection, topic) of the subject program is also necessary for the further formation of parallel variants of test. To use one variant of test for the audience is impossible in terms of objectiveness: there are cheating, prompting, and incorrect behavior and so on. That's why the developer of tests (or organizer of the test development) has to create several variants so that respondents who are tested were in the same conditions: one or two variant for one class or several dozens for federal or regional testing. They become parallel if the test items in them are related to the same elements of program content and control the mastering of the same knowledge and skills. As technological bases of forming parallel variants of the test it is better to use the bank of test items (Biberina, 2000).

Drawing up such documents as test specification and test plan with which the stage of planning a test is over take place before developing parallel variants of the test.

The fulfilment of specific test items significantly depends on their form: on how the question is formulated or how the test item looks like. So the evaluation of the academic achievements on the results of the test is mainly determined by the professionalism of the developer, how well he knows the peculiar features and the possibilities of the different forms of test items.

## 5. Conclusion

The conducted research in accordance with the set aims and objectives allows to formulate the following findings.

1. There are several classifications of the test item forms in the national and foreign literature.

In spite of differences in classification the following test item forms might be pointed out:

Test items with a choice of one or several answers from the given variants.

Test items with multiple choices and matching.

Test items with alternative answers.

Test items with the free constructed answer:

– Test items with the short specified answer (additions in the type of unfinished sentence or sentences with the omitted word, or the question, which suggests the answer in the form of one or two words).

– Test items with full answer.

The given classification considers not only the form of information delivery, but also suggested algorithm of the students' activity on the information processing.

All variants of the test item forms are numerously described in the special literature, however, the authors of many brochures with the tests and monitoring and measuring materials usually break methodological principles, regulations and norms, using the test items created with the violations. This dictates the necessity to consider the methods and the examples of the developing all possible test item forms with regard to monitoring historical knowledge and skills and to analyze their eligibility to be used in different types of pedagogical monitoring.

2. For each content element determined to be included in the test, it is necessary to set the level of mastering and type of knowledge and skills which the student has to obtain during the education. This is specifically

important moment in planning a test from the point of view of activity and competence-based approaches in education. Otherwise the education will be mainly oriented to form the knowledge of first level (reproduction of the content by memory) and only factual knowledge will be monitored.

3. While developing the test it is essential to remember that the result of the history education must be mastering specific types of activities for this subject and general educational activities related to different levels of mastering academic material. Test items must control all these types of activities in maximum amount. That's why the developer needs in their detailed list.

4. The clarification of the type of knowledge and skills which are formed while learning this or that section (subsection, topic) of the subject program is also necessary for the further formation of parallel variants of test. To use one variant of test for the audience is impossible in terms of objectiveness: there are cheating, prompting, and incorrect behavior and so on. That's why the developer of tests (or organizer of the test development) has to create several variants so that respondents who are tested were in the same conditions: one or two variant for one class or several dozens for federal or regional testing. They become parallel if the test items in them are related to the same elements of program content and control the mastering of the same knowledge and skills. As technological bases of forming parallel variants of the test it is better to use the bank of test items

5. So to develop the quality test the selection (development) of test item might have several iteration procedures, unless there would be the correct proportion in the reflection of the subject content and the types of the monitored activity of the student by the test.

## 6. Recommendations

The practical significance of the research is that the results and findings might be used by the history teachers in the process of developing and using the tests for all forms of control, independent construction of tests for everyday activity, analysis of testing results and inspection of the content of prepared test items.

## Acknowledgement

The work is performed according to the Russian Government Program of Competitive Growth of Kazan Federal University.

## References

- Andersen, P., & Morgan, J. (2011). *Razrabotka testov i anket dlya natsional'noi otsenki uchebnyh dostizheniy*. Moscow.
- Avanesov, V. S. (1998). *Compozitsiya testovyh zadaniy*. Moscow.
- Avanesov, V. S. (2005). *Forma testovyh zadaniy*. Moscow.
- Biberina, A. V. (2000). *Testy kak sredstvo proverki urovnya podgotovki uchastivshisya po istoriyi drevnego mira i srednih vekov* (Unpublished master's thesis). Moscow.
- Chelyshkova, M. B. (1995). *Razrabotka pedagogicheskikh testov na osnove sovremennyh matematicheskikh modelei*. Moscow.
- Chelyshkova, M. B. (2001). *Teoriya i praktika konstruirovaniya pedagogicheskikh testov*. Moscow.
- Chelyshkova, M. B., Zvonnikov, V. I., & Davydova, O. V. (2011). *Otsenivanie kompetentsii v obrazovanii*. Moscow. 2011.
- Efremova, N. F. (2003). *Sovremennye testovye tehnologii v obrazovanii*. Moscow.
- Efremova, N. F. (2010). *Formirovanie i otsenivanie kompetentsiy v obrazovanii*. Rostov na Donu.
- Emelyanov, A. V., & Yatsenko, M. P. (2002). Otsenka urovnya podgotovlennosti uchastivshisya v uchebnom istoricheskom poznanii. *Shkol'nye tehnologii*, 6, 189-193.
- Ivanenko, N. A., Burdukovskaya, E. A., Yunusova, A. B., Mukhametshina, R. F., Letyaev, V. A., Islamova, L. R., & Yeremeyeva, T. S. (2015). Corporativity as a Condition for Developing Teaching Staff. *Review of European Studies*, 7(4), 49-53. <http://dx.doi.org/10.5539/res.v7n4p49>
- Kadnevskiy, V. M. (2004). *Istoriya testov*. Moscow.
- Khanmurzina, R. R., Sinitzyn, O. V., Kaviyev, A. F., Burdukovskaya, E. A., & Lebedeva, I. V. (2015). Practical Recommendations for Optimizing the Process of Formation and Development of Academic Mobility of Students in Higher Educational Institutions of Russia. *Review of European Studies*, 7(1), 41-45.



- Khayrutdinov, R. R. (2014). Museum Conference of 1926 in Kazan. *World Applied Sciences Journal*, 31(2), 193-195.
- Lerner, I. Ya. (1982). *Razvitie myshleniya uchschihsya v protsesse obucheniya istorii*. Moscow.
- Levina, E. Y., Yunusova, A. B., Fayzullina, A. R., Rassadin, S. V., Lotfullin, M. R., Nuriyeva, E. N., & Vlasova, V. K. (2015). Federal Public Administration Implementation in Vocational Education. *Review of European Studies*, 7(4), 98-104. <http://dx.doi.org/10.5539/res.v7n4p98>
- Mayorov, A. N. (2000). *Teoriya i praktika sozdaniya testov dlya sistemy obrazovaniya*. Moscow.
- Mikhailychev, E. A. (2002). Sistema pedagogicheskoi diafnostiki: klyuchevye ponyatiya i principy. *Pedagogicheskaya diagnostica*, 1, 46-76.
- Mikhailychev, E. A. (2003). Metodiki didakticheskoi diagnostiki. *Shkol'niye tehnologii*, 5, 97-101.
- Mikhaleva, T. G., & Khlebnikov, V. A. (2000). Metodicheslie printsipy formirovaniya edinogo banka testovyh zadaniy *Tezisy dokladov Vtoroi Vserossyiskoi konferentsyi "Razvitie sistemy testirovaniya v Rossyi"*. Chast' 3. Moscow: Prometei.
- Pereverzev, V. Yu. (2003). *Criteria'no-orientirovanoe pedagogicheskoe testirovanie*. Moscow.
- Pereverzev, V. Yu. (2005). *Tehnologiya razrabotki testovyh zadaniy: spravochnoe rukovodstvo*. Moscow.

### Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).