

DEVELOPMENT OF EDUCATIONAL MEASURING TESTS FOR INTERNET-EXAM IN HIGHER EDUCATION OF RUSSIAN FEDERATION

V. Navodnov, A. Maslennikov, V. Kiseleva and L. Sharafutdinova
National Accreditation Agency of the Russian Federation (NAA), Russia

Abstract

The testing for assessment of students learning outcomes in Russia carried out by means of various tests and in different periods of the academic year do not allow to adequately compare the results either with other HEIs or with the requirements of State Educational Standards (SES).

The article looks at possible approaches to solving the problems in the framework of the Internet-exam conducted by the National Accreditation Agency of the Russian Federation. The use of tests from the uniform federal database and implementation of the uniform technology for analysis of results provides for openness and transparency of evaluation procedures. The involvement of HEIs' academic community in the process of elaboration, up-dating and wide discussion of testing materials is an important aspect of this initiative. It corresponds to the requirement of independency of evaluation procedures both from education authorities and HEIs.

The assessment of content and level of mastery of individual disciplines is made by specially developed tests containing tasks which are developed in full compliance with the requirements of SES. The creation of testing materials required thorough analysis of standards made for each discipline in the aggregate of higher educational programs. Based on the principle of invariance of discipline content for groups of educational programs a new approach to the development of testing materials framework was suggested to supervise the implementation of standards' requirements. This also allowed to reduce costs and efforts in developing indicators for each discipline.

The top-priority task of contemporary education is the development of education quality management system. The main constituent of the system of student's training quality management is the evaluation of content and level of

students' training in terms of compliance with the requirements of state educational standards.

In some universities of the Russian Federation there are quality management systems the main function of which is monitoring students' training level. However the evaluation of students' training is carried out locally and excludes the possibility of comparing the results of the training. Internet-exam in professional education conducted by National Accreditation Agency of education helps to develop the unified system of evaluation students' training level.

A significant feature of students' training evaluation level of Internet-exam is full content coverage within the State Education Standards. For these purposes a set of specified Evaluating Educational Measuring Tests (EEMT), which sufficiently correspond to the level and content of Education Standards requirements was developed. The principal structure component of the measuring instrument is the section of the discipline (didactical unit of the content). Each section is presented by tasks on several sections. Such approach to development of Educational Measuring Tests can provide for high content validity of the measuring instrument in terms of SES. The evaluation of Internet-Exam results is done using the methods of acquiring the didactical units of the discipline.

The development of Evaluating Educational Measuring Tests for Internet-Exam consists of several stages. First stage is devoted to the macroanalysis of SES contents for the whole of professional training directions (basic academic programs). The results of discipline content analysis are presented in the form of the taxonomical chart [1]. For example, the SES content analysis for «Mathematics» proved that for 600 academic programs realized in Higher Education Institutions of Russia, it is quite enough to select 18 basic didactical units of the discipline contents. The set of didactical units of the discipline content for a group of academic programs is content invariance.

The proposed method of invariants pickup from the content of Educational Standards allows to minimize the quantity of Educational Measuring Test structures. For example, for 600 basic academic programs, only 34 structures of measuring materials of the discipline "Mathematics" are required. Thus, using this new approach to measuring instrument development it is enough to work out topical didactical units of ES content and there is no necessity to develop separate measuring instruments for every academic program.

In the process of EEMT topical structure development the amount of time required for the discipline study is analyzed. On the basis of time analysis the complexity level for each kind of training is determined. Significant differences in time required for similar disciplines allow to study these disciplines at different level of profoundness. Therefore different complexity level tasks are required to check content and level of students' training. As a rule, in the process of EEMT development for Internet-Exam two or three level tasks are used.

At the next stage the descriptors of EEMT topical tasks are developed. They include the list of didactical units to be checked, the level of complexity and other expert characteristics. On the basis of descriptors the sets of academically equivalent tasks in one topic are developed. It is essentially important in test development to make distractors taking into account students' typical mistakes. This approach allows to analyze students' level of training more thoroughly and to identify their obstacles.

The third stage implies the procedure of approbation, task quality analysis and their revision. The following statistical features of the tasks are to be pointed out:

- *Task resolution ratio;*
- *Task differentiating feature (discriminatory power);*
- *Point biserial correlation ratio.*

Task resolution ratio is defined by the number of students, who coped with the task. The most effective is the task with resolution ratio above 0.6.

Task differentiating feature (or the task selectiveness ratio) allows to divide the students according to the level of mastering the subject. Differentiating feature is defined as the difference of task resolution ratio for high level student group and for lower level student group. The higher task differentiating capacity, the more valid and reliable is the measuring instrument.

The point biserial correlation ratio allows to analyze the connections between a single item and the whole test performance. The positive correlation for the right answer shows that the students with high test mark cope with the task to be analyzed. The tasks with the point biserial correlation ratio over 0.3 are included into task bank. The distractors must have negative point biserial correlation ratio, over 0.2 in absolute value.

The analysis of task statistical characteristics is the basis for task bank updating in disciplines. At the moment there are 43 EEMT banks for three groups of higher education disciplines and 10 banks for secondary vocation disciplines.

The forth level of test bank development is reviewing and certification.

One of the ways to make the evaluation of students' training level more reliable is to offer a wider range of test disciplines. It became possible due to multidisciplinary approach and selective test method. Multidisciplinary testing is represented as students' simultaneous testing in several disciplines of the same group of disciplines. Thus the procedure of evaluating the level of both a separate discipline and a group of disciplines acquiring became possible.

The method of multidisciplinary EEMT development is also based on structural approach of the ES content analysis.

Bibliography:

1. Maslennikov A.S., Kiseleva V.P. Methodology for developing content invariance of the discipline "Mathematics" for EEMT development. 8th Vavilov's readings. The outlook of the modern society in the focus of scientific knowledge and practice: Collected articles / Edited by professor Shalaev V.P. In 2 volumes. – Yoshkar-Ola: MarSTU, 2004.-V.2 – p.94-96.
2. Navodnov V.G. Mathematical models SAP EMM. Preprint №4/97. Yoshkar-Ola: MarSTU, 1997. – 72p.
3. Maslennikov A.S., Saveljev B.A. The evaluation of student's comprehension level aimed at certification of vocational education institution: Text book. – M.: Logos, 2003. – 136 p
4. Navodnov V.G., Maslennikov A.S., Kiseleva V.P., Sharafutdinova L.N. The application of Internet-Exam in the field of professional education for quality evaluation of students' training. Contemporary diagnostic means for evaluation meant for education quality certification and the use of IT: X1 symposium materials «Qualimetry in education: methodology, methods, practice ». Part3. – M.: Research center of specialists' quality training problems, 2006. – p.48-54.
5. Crocker, Linda and Algina, James. Introduction to classical and modern test theory. Florida. – 1986. – 527p.