World Applied Sciences Journal 30 (11): 1610-1617, 2014
ISSN 1818-4952

DOI: 10.5829/idosi.wasj.2014.30.11.14222

Testing as the Assessment Tool in the Engineering Higher Education: Positive and Negative

Marina Prakhova, Gennady Kolovertnov and Ernst Shalovnikov

Ufa State Oil Technical University, Ufa, Russia

Abstract: The article examines the role and place of testing in the evaluation of knowledge in engineering higher education. The urgency of this issue is connected with Russia's accession to the Bologna agreement and the transition to the new generation of educational standards, implementing competence-based approach to learning. The analysis of the advantages and disadvantages of testing has done. The paper presents also the results of the authors interviewing the graduating class of pupils and students of technical university. Some recommendations on the use of testing in higher school are provided.

Key words: Reform of higher education · Educational supervision · Testing · Interviewing · Competency approach · Domain-specific knowledge

INTRODUCTION

Reform of higher education, the ongoing in Russia, led to a sharp increase in the quality requirements for the training of professionals of all fields.

The transition to two-tier system of education bachelor-master is the result of Russia's accession to the Bologna Convention in 2003. It brought many changes in higher education. They are aimed at its modernization. Her main areas:

- To bring direction and content of training in line with modern science and technology development that led to the emergence of new federal state educational standards of the third generation;
- Establishing the content and technology education needs cotemporary society (the transition from reproductive learning, i.e. the transfer of "ready-made of knowledge," to the competence learning, the purpose of which - the formation of professionally significant skills of students, the so-called competence);
- Development of new information technology training;
- The creation of an objective and effective quality control system received education that meets international standards.

Quality education is the main criterion for evaluation of an educational institution. There are many methods for achieving this goal. One of them-the organization of effective pedagogical control.

One way of such control is testing. The question of testing as a tool to assess knowledge in higher education is one of the most controversial and is a lively debate in almost any university. This is true not only for Russia but also for other countries. Both supporters and opponents of the test result valid arguments in support of its position [1-5].

Recently, interest in this issue again brightened. The reason was the proposal to introduce an independent testing for all graduates of bachelor in the country. This test should replace the existing state exam. It’s goal - to promote the mobility of students within the country and become an indicator of the quality of bachelor in higher schools. The Higher School of Economics is preparing a pilot project of this exam. This is supposed to be a standardized exam, for example, as GRE subject test. This exam is for bachelor graduates wishing to enter the U.S. master's programs, it includes a test of language proficiency and knowledge in mathematics.

Nevertheless, the ratio of the testing is very ambiguous and this applies to all communities connected with higher education in one way or another - teachers, employers and students. In this case the profile of

Corresponding Author: Prakhova, Ufa State Oil Technical University, 1, Kosmonavtov street, Ufa, Russia, 450062.

education (humanitarian or technical) is irrelevant: the supporters and opponents of the tests are in all universities.

Methodology and Research Goals: The current tendency to make testing as the main tool for the control of knowledge, including for the final assessment, was the reason for this article.

We are aware that, strictly speaking, the choice in favor of testing is already made by external factors (included in the Bologna Convention, the implementation of quality management systems, abandonment of traditional examination sessions in favor of the point-rating system (PRS), etc.).

So the main goal of our research is to analyze the pros and cons of testing in terms of assessing the quality of higher education.

This analysis will help to solve the following problems:

- To maximize the advantages of testing and to minimize the impact of it's shortcomings;
- To find any niche in the kinds of the assimilation of discipline control, where the use of tests is the most appropriate.

Have been used research methods such as descriptive method, analytical review and the standard survey method.

The survey was conducted in the following focus groups:

- Students of the graduating class of the Lyceum;
- Students of oil technical university (full-time education);
- Students of oil technical university (distance learning).

RESULTS AND DISCUSSION

Currently, there are three main areas of testing:

- Education (in connection with increasing duration and complexity of training curricula). The evaluation function of testing is double: first, it allows you to assess the knowledge of a discipline specific training and secondly - the correctness of the curriculum subjects, the choice of educational methods, etc. factors;
- Training and selection (due to the increase in the rate of growth and complexity of production). At least testing provides an opportunity to identify people who obviously are not suitable for any professional activity and to determine the recommended area;
- Psychological counseling (in connection with the acceleration of social processes). Testing can identify with a certain probability the actual level of development of human skills, knowledge, personal characteristics, etc.

Science knowledge assessment by means of tests began in the late XIX-early XX century, at the intersection of psychology, sociology, pedagogy and other so-called behavioral sciences. Tests themselves have appeared much earlier. In ancient China they were used for the selection of individuals for public office [6].

The ancestor of the tests in its present form is considered to be an Englishman, Francis Galton. He introduced three fundamental principles to the theory of testing which are used till this day:

- Using the same series of tests to a large number of subjects;
- The statistical treatment of results;
- The selection of standards assessment.

While in the West testing used as a tool for evaluation of a quality for over a century, in Russia, his fate was thorny and unpredictable.

In the early twentieth century, the post-revolutionary period, when any of the methods of control of knowledge have been canceled, the use of tests was banned [7].

Then in the 20s, the education revolution occurred and many Soviet scientists and educators engaged in research relating to all aspects of testing. But in the early 30-ies of XX century, all these works have been criticized and then, in 1936, were banned altogether.

Centre for Development testing has shifted to the West, where investments in the development and evaluation of tests increased year by year.

Tests used to solve for admission to universities, a prestigious job and to address other vital issues were called «high-stakes tests».

However, this led to another extreme.

Evaluation of the accuracy of measurement results and the adequacy of the interpretation of the testing results were not supported by a serious check and the testing procedure often was not standardized. Therefore,

expansion of testing for the formation of individual life trajectories strongly exacerbated the problem of the influence of the testing results on the social, psychological and economic aspects of life.

Not surprisingly, the massive use of imperfect tests, whose results are affected by certain decisions, leading to performances against tests, which often involved students.

In the Soviet Union resumption of testing technology has coincided with the beginning of the introduction of computers that allowed to formalize the process of testing and facilitate statistical processing of large volumes of data.

The first official recognition of testing as a way to test the level of knowledge appeared in the December 1988, initially as entrance tests, including technical schools and since 1995 - as a way to final certification of schoolchildren [8].

It should be noted that all matters relating to testing, are involved in testology science. This is an interdisciplinary science to create high-quality and evidence-based diagnostic techniques for measuring. In fact, testing is a mathematical technique to create and use a variety of tests (psychological, technical, medical, professional, educational, etc.) and is a branch of applied mathematics and cybernetics.

To date, the global leader in the use of test technologies in education are the USA (according to other reports [8] Testing Centre are the Netherlands, followed by Japan, Australia, USA). All scientific and methodical literature published in English. In the West the rating of the university depends on the level of test control: the higher the level of testing, the higher the rating.

At the same time the Americans themselves say that although the national standardized tests are dealt with each year better, the gap between the level of knowledge of American students and their peers from European and Asian competitors increases and among the reasons for this is called “the holy faith in the standard test” [9]. Not less eloquently the editorial in the newspaper “Los-Angeles Times” called its article: “Testing kills learning”.

In higher educational institutions of Russia the testing so far belongs to the innovative methods. This is due to the fact that most teachers haven’t got scientific knowledge of the methodology and technology of testing.

Creating a reliable, valid test with stable characteristics is very complex and time-consuming process.

In order to test really was an instrument of objective assessment of knowledge of discipline, we must carefully consider all its components - the actual test items, the testing procedure and the model aggregation (how to interpret the results).

As a result, most tests are designed for teachers of their own risk. Of course, this is still a positive thing and should be encouraged. These tests are basically a set of 20-30 questions, each of which is invented by a few answers. Their degree of plausibility depends on the teacher’s time and his imagination. The practical value of such tests is highly questionable.

The authors of most publications on testing, for example [10-12], point the following features as the undoubted benefits of testing:

- Control objective, as all students are on an equal footing in the process of the testing and processing the results of the same criteria, subjectivity of the teacher is excluded;
- Efficiency of the process and results (especially if the computer test is used), which makes testing extremely useful tool for monitoring the quality of the educational process within the quality management system;
- Standardization of test items and test procedure allow to evaluate the level of knowledge in any discipline of students from various universities (on-line testing) and then use these results for the accreditation of specialties;
- Analysis of test results can be used to make appropriate management decisions regarding the correction of curriculum subjects, choosing educational technologies, methodological support and discipline, etc.;
- Test lets you check the knowledge for the entire discipline, while all exams are a kind of lottery: what questions will fall on the ticket. Even if the teacher has a habit to ask supplementary questions, a controlled amount of knowledge on the exam is much smaller;
- Provides a more accurate assessment of knowledge, because scale evaluation of the test has a lot more gradations (in accordance with the number of jobs) than the usual scale of assessment of knowledge;
- Testing requires far less expensive to conduct and inspection, than examination (excluding the cost of development of test items themselves, that, ideally, is done once).

Disadvantages of testing (some of them also reported in the cited works), in our opinion, confirm the thesis that at least part of them is the obvious extension of its advantages:

- Development of quality test questions is a lengthy and laborious process, but the opinion that the test is created once and for all, is profoundly mistaken. Practice shows that the variants of test items should be changed frequently, because they soon become known "to the public." When using computer-based testing preferable to have a relatively bank of jobs that are completed in the options at random. Change, such jobs can be much less, but their initial amount increases significantly. There are a number of disciplines (mainly technical), which are significantly updated every year, so this implies a change in the tests;
- Declared breadth of material in the test is clearly the opposite direction: although it requires the student familiarity with all the stuff, but very superficial, at the level of concepts, not fact;
- In any testing (as well as on the exam) is an element of chance; it can be due to random error in the design of the response ("crossed out the wrong cell") or, that happens more often, guess the answer. It distorts the results of the test and leads to the need to use fairly complicated mathematical apparatus to handle them, which allows to take into account the probabilistic component;
- Even the impeccably compiled test items do not allow us to estimate non-standard, unconventional thinking, although it is desirable to identify such students as early as possible - they are subsequently Masters and postgraduates. This shortcoming, in our opinion, is the most important - tests do not allow students to demonstrate their ability to think.

Quality education is determined equally by those who give knowledge and those who receive them. To study the views of students on tests, we conducted a survey.

The questionnaire was entitled "Test as a way to assess knowledge: the pros and cons" and contained the following questions:

- What do you prefer-a test with multiple choice questions or colloquium, jobs which require a detailed reasoned response?
- Does it matter to you, how many correct answers can be in the test task and the order in which questions are asked (on the test sequence or at random)?
- Do you usually answer the test questions intelligently or more are relying on luck?
- How do you think, the score for the test matches your knowledge is too high or too low?
- Your personal opinion about the tests as a way of assessing knowledge (strengths and weaknesses).

The questionnaire results are somewhat expected, but something quite unexpected.

Answers to the first four questions are presented as charts 1-4. Designation focus groups corresponds to the following groups:

- Pupils of the graduating class of the Lyceum;
- Students of Ufa State Oil Technical University (full-time education);
- Students of Ufa State Oil Technical University (distance learning).

According to the results of the questionnaire, preferences of school pupils and students are opposed.

100% pupils prefer the test and half of them rely on luck. It's interesting that in this approach the structure of test task however is not indifferent to them. Apparently, if test items are ordered, the answer is easier to guess. Evaluation is often overestimated (75%).

Full-time students (we interviewed students of two training profiles: "Automation of technological processes and production" and "Design and operation of oil and gas pipelines, gas storage facilities and tank farms") in the majority are opposed to testing (testing was preferred only 28% of respondents). Furthermore, even these 28% of respondents when answering the last question, noted as the main drawbacks of testing the next ones: a high probability of random guessing the correct answer and the inability to evaluate the student's ability to reason logically and to prove their point. Here are two of the most characteristic responses to this question: "Tests are the most favorite way to assess students' knowledge, because you can guess the correct answer or write off a neighbor, but do not fully reflect the student's knowledge on the subject" and "The test does not show the thought process which is valuable in itself. So generally a test is for the lucky students and a colloquium - for thinking ones!".

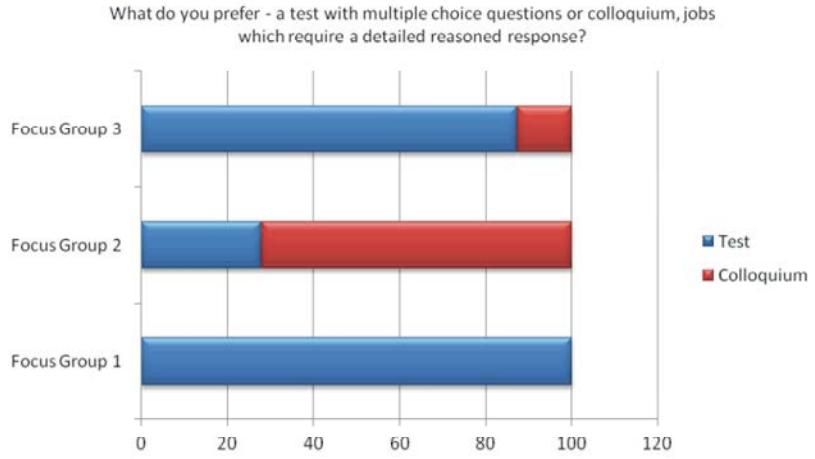

Chart 1: Distribution of answers to question 1

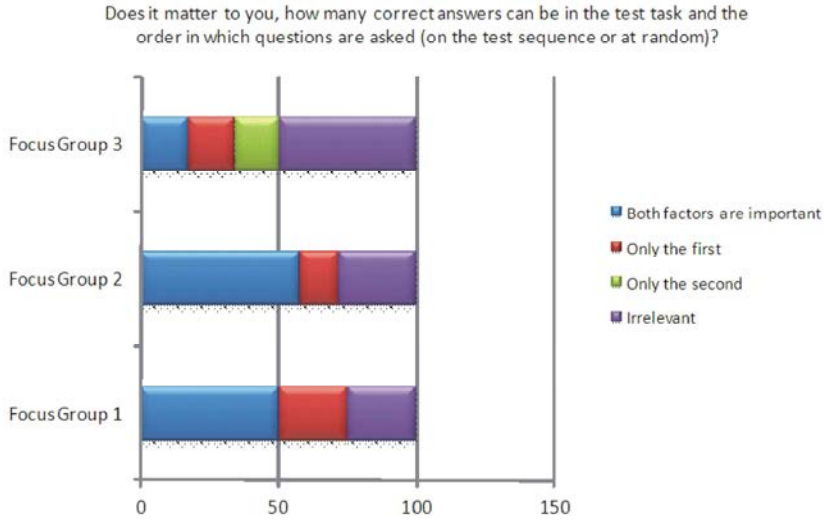

Chart 2: Distribution of answers to question 2

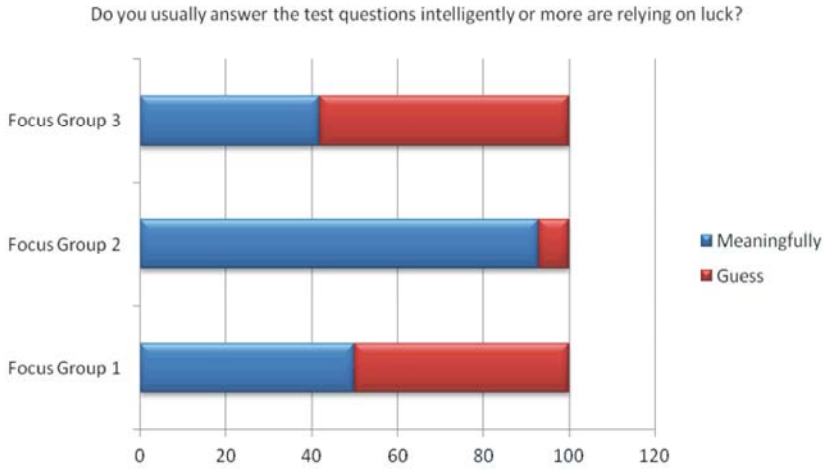

Chart 3: Distribution of answers to question 3

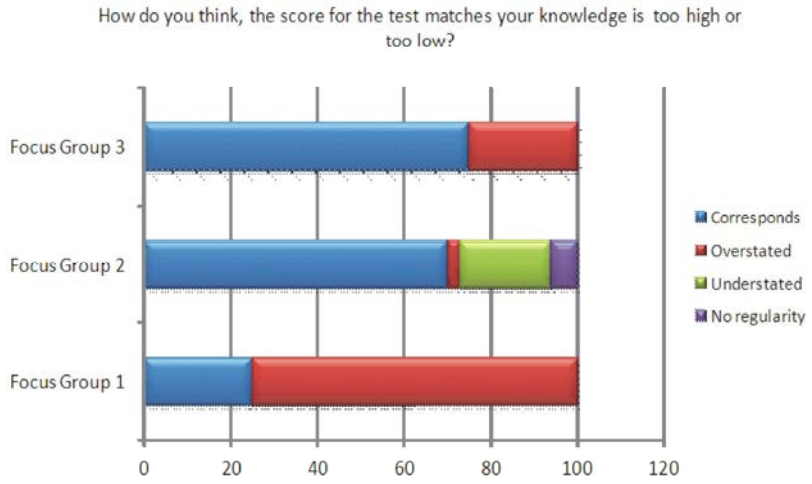

Chart 4: Distribution of answers to question 4

As for two requirements for the test tasks (only one correct answer and the location of jobs in the sequence corresponding to the sequence of discipline material), for the majority of respondents it's important.

Among distance learning students the picture is reversed. The test was chosen by 66% of respondents and 58% of respondents answer at random. Hence, the expected high degree of indifference to the structure of the test questions - for half of the respondents, it does not matter.

Thus, the students themselves note as the main advantages of the test, small amount of time (it is noted by almost all respondents) and convenience (to use ready-made wording to the answer) and as deficiencies - inability to assess the ability to justify and defend their opinions, as well as the frequent discrepancy between the results of real knowledge. As for the objectivity of testing, many respondents emphasize that the test is objective insofar as exempt from *the teacher's subjectivity*, but an *objective assessment of knowledge* isn't.

All these arguments apply to models of education, in which the quality of it is estimated by the quantity of *acquired knowledge*. The authors of [8] aim to "…persuade opponents to promising test technology for assessment of the quality of students' knowledge". Indeed, in many ways they are convincing. But not for nothing they use exactly this semantic concept-"*knowledge*".

However, the concept of knowledge has several varieties. The authors of [13] divide domain-specific knowledge into four types—declarative ("knowing that"), procedural ("knowing how"), schematic ("knowing why") and strategic ("knowing when certain knowledge applies, where it applies and how it applies").

In our view, the transition to new standards, implementing competence-based approach to education, makes testing in its current form, unsuitable for assessing the quality of education tools. The main reason is that testing evaluates only declarative and procedural knowledge, partly schematic knowledge and never strategic ones. This is especially true of domain-specific knowledge.

Competence is an integral concept. It is the behavioral model of professional and social activity, the ability to establish and implement a graduate of the relationship between "knowledge - the ability" and the specific situation.

It consists of three components.

First, this knowledge is not just a luggage information. It is mobile knowledge that can quickly find the information, to weed out unnecessary and to translate the experience into its own activities. We can say that this knowledge is not only and not on a particular subject, but about how, at any time to find another appropriate knowledge.

Second, this is the ability to use knowledge in a particular situation, it is an understanding of what knowledge is lacking for this situation and how to get it.

Third, this is an adequate assessment-myself, the world, my place in the world, specific knowledge of the necessity or uselessness of something for any activities.

It's unable to determine the exact contribution of any discipline in the formation of any competence, so the amount of test results for individual subjects may not be assessing the quality of graduates.

In addition, the use of testing is complicated by the nonlinear development of the content of the educational program [14], consisting of a large number of disciplines for the selection and individual training plans.

CONCLUSIONS AND RECOMMENDATIONS

In this situation, the test has two alternatives. The first one is to use it as a "niche" tool for some particular situations (e.g., blitz-interview with the purpose of admission to the laboratory work, the input control on discipline, self-study test subjects, etc.). Development of these test items do not require special training of teachers, saving the time it takes to control, i.e., the benefits of testing implemented and cons remain "behind the scenes".

The second alternative on the contrary, is the yield of testing for a higher level. For this purpose, the creation of a new generation of tests is necessary. It has to be integrated tests (not the same with interdisciplinary!), which can evaluate the level of training in the direction of a whole training profile or separately for cycles directions. As a benchmark for these tests, you can use the methodology and practice of professional tests, developed by employers' associations. Tasks of such tests (at least some of them) must be open, i.e. requiring a free text response with the necessary justification. This will simulate a simple situation, which the student can meet at work.

Of course, such task will require some more time to perform as well as for testing, but benefit will be much more.

Naturally, the transition to this level of testing requires a completely different attitude to the development of test items.

It is only possible if at each department it will be at least one or two teachers specially trained in the development of tests, examination, testing techniques. They need to educate other teachers.

And, of course, such testing should be performed only by means of special computer programs.

But even in this case, test will not become a comprehensive tool for the knowledge control. In accordance with the principles of the formation of competences it's necessary to develop the communicative abilities of the student, he must know not only the right, in his opinion, the answer, but to be able to explain why it is loyal, to be able to defend his position. You can learn this only through verbal communication between student and teacher, discussion in seminars. These forms of the educational process have always been important and will remain so, despite the change in the model of higher education. A testing must to "know its place" and used as a form of partial control (e.g., landmark) of the modern educational process.

REFERENCES

1. Gajiæ, O., S. Markov, B. Raduloviæ and M. Stojanoviæ, 2012. The Analysis of Students' Opinions on the Role of Knowledge Tests in Higher Education // The New Educational Review, 30(4): 114-125.
2. Popham, W. James, 1999. Why Standardized Tests Don't Measure Educational Quality // Educational Leadership, 56(6): 8-15.
3. Gibbs, G., 1999. Using assessment strategically to change the way students learn //Assessment Matters in Higher Education: Choosing and Using Diverse Approaches. Buckingham: Society for Research into Higher Education.
4. Gow, L. and D. Kember, 1993. Conceptions of teaching and their relation to student learning //British Journal of Educational Psychology, 63: 20-33.
5. Assessing Learning in Australian Universities: Ideas, strategies and resources for quality in student assessment. [Electronic resource]. URL www.cshe.unimelb.edu.au/assessinglearning.
6. Kim, V., 2007. Testing academic achievement. Monograph. Ussuriysk: Publisher UGPI.
7. Maltsev, A. and A. Naymushina, 2008. Testers in education: yesterday, today and tomorrow // Problems of Education, Science and Culture, 60(24).
8. Bulanova-Toporkova, M., 2002. Pedagogy and Psychology of Higher Education: Handbook. Rostov-on-Don: Phoenix.
9. Kirillova, S., TIMSS on both your houses! [Electronic resource]. URL http://ps.1september.ru/ articlef. php?ID=200501006.
10. Kosukhin, B., G. Loginova and Y. Loginova, 2008. The role and place of testing in high school. Higher education in Russia, 1: 94-97.

11. Test method for assessing students SSU, “for” and “cons”, 2009. [Electronic resource]. URL http://www.universite.ru/index.php.
12. Vasiliev, Y. and V. Gordin, 2010. Experience testing in St. Petersburg branch of state-owned University. Higher School of Economics: results, problems and prospects [Electronic resource]. URL http://www.hse.ru/data/2010/04/30/1216579667.
13. Shavelson Richard, J. and L. Huang, 2003. Responding Responsibly to the Frenzy to Assess Learning in Higher Education //Change, 1(2).
14. Roger, B., 2012. The Seven Red Herrings About Standardized Assessments in Higher Education (Foreword by P. Ewell, Commentaries: M.A. Miller, T.L. Rhodes, T.W. Banta and G.R. Pike) // National Institute for Learning Outcomes Assessment, September.