

Mathematical Model and Qualimetric Assessment of Graduate Education Quality in Environment Saturated with Information and Communication Technologies

Ernst Raisovich Safargaliev¹, Irina Ilyinichna Eremina², Savitsky Sergey Konstantinovich¹ & Valentina Aleksandrovna Camelina³

¹ Kazan (Volga Region) Federal University, Branch in Elabuga, Russia

² Kazan (Volga Region) Federal University, Branch in Naberezhnye Chelny, Russia

³ Mari State University, Russia

Correspondence: Ernst Raisovich Safargaliev, Kazan (Volga Region) Federal University, Branch in Elabuga, 89 Kazanskaya Street, Elabuga, Republic of Tatarstan, 423600, Russia. Tel: 7-937-291-5657. E-mail: safargaliyeva@mail.ru

Received: October 5, 2014 Accepted: November 7, 2014 Online Published: January 27, 2015

doi:10.5539/ies.v8n2p78

URL: <http://dx.doi.org/10.5539/ies.v8n2p78>

Abstract

Our model uses a complex quality assessment method of academic training of graduates. We have used a synthesized model, which is based on both positive and negative experience and uses qualimetric approach in quality assessment. The complex factor of the graduate's competency is a sum of scores for knowledge, skills and expertise in subjects that comprise a compulsory minimum of the basic graduate educational program under the existing state educational standard. The scores for knowledge, skills and expertise in a subject are given to a student after tests.

Keywords: education information system, future IT professional, graduate competency, mathematical model, experiment interpretation, testing

1. Introduction

The problems of higher education quality management are tackled in Russia in different directions: development of theory of educational quality management (educational qualityology); standardisation (state educational standards and quality management systems based on international standards); establishing and developing state and public quality management mechanisms (attestation, licensing, accreditation); theoretical and methodological research with regard to the assessment of the educational status and its development trends. The priority in developing quality systems in higher education should lie with the objective assessment that provides a scientific base for analysing the results, for functioning and development of educational quality management systems.

Various models are used for assessment of the quality of students' education. The assessment is done in Russian higher education institutions as part of a complex audit. The models are also used when taking a decision about accreditation. The decisions based on the information gathered from different models can differ significantly.

Measurement in qualimetry is a unique quantitative statement of quality in the units and scale of the chosen measurement scale. As a result of this measurement, a quantitative unit of quality expressed by a quality score is defined. Quality scores of students' education are the results of many tasks performed by them and meant to detect their competencies and skills in different situations.

Currently, there exist several models to assess both expertise, knowledge, skills and competencies. Clearly, they all have their drawbacks. Some of them are:

- Threshold value fluctuation;
- Two quality levels only ("passed" or "failed");
- Different number of tasks in sections not being recognized;
- Task difficulty levels are not taken into account;

- Small number of test questions brings about noncompliance of the student's education level with the state educational standard.

Having analysed the above drawbacks, we have come to the conclusion that it was necessary to use a synthesised model, which would be based on both positive and negative experience and would use qualimetric approach in quality assessment. Having said that, we need to define a quality benchmark, the ways to measure certain characteristics that form quality. We also need to define scores to be included in the content, the ways to assess them. It is essential for the model of a complex quality score to recognize the significance and allowable change limits of certain characteristics to the degree that would conform to the real conditions of professional activity.

Besides, Federal State Educational Standard of Higher Vocational Education of the third generation requires that educational policies and practices of all higher education institutions should now be built on the competency building approach. It expands the definition of students' academic training quality: now it is the whole of knowledge, expertise, skills and competencies.

2. Method

In our model we use a complex method of quality level assessment. Complex score of the graduate's competency is defined as the whole of scores for knowledge, expertise, skills and competencies in the subjects that compose the compulsory minimum of the basic educational programme of the graduate's training under the existing Federal State Educational Standard.

The effective tool of an independent assessment is the educational environment. One of the most well-known and popular e-learning control systems capable of developing professional and information and communication competencies provided there is right methodological support is LMS Moodle (Modular Object-Oriented Dynamic Learning Environment). In Kazan Federal University (KFU), the educational environment based on LMS Moodle has been in place for a few years; thus, certain experience of working with the above as well as observations and conclusions about the influence of the educational environment on the academic process and its active participants, i.e. students and teachers, have been gathered. The main means of testing the level of professional and information and communication competencies is the final test in each suggested section and case studies.

Statistics system is based on Russian metric system that enables us to talk about the quality of pedagogical measurements. However, the problem is the complexity of statistical processing of test results, calculation and further interpretation of the required characteristics.

3. Results

Based on the analysis the following conclusions can be made:

- for those tests that are built on elementary ("easy") questions, well-known methods that increase reliability do not produce expected results;
- small quantity of tasks in a library of questions and in a test makes the test quality assessment unreliable and the application of improvement techniques impossible.

Therefore, the system for analysing test statistic results within an academic course, which has been described above and implemented in Moodle network control system, can be an effective tool for the managers in controlling the quality of test materials that are designed by a teacher to diagnose professional competencies and communication and information competencies; it can also be a handy tool for the teacher in improving a test and test tasks to increase the education quality and to control the academic progress.

To perform quantitative assessment of professional and communication and information competencies, the qualimetric analysis has been used. It defines the integrated coefficient (K) for the development level of the addressed competency of future IT professionals.

The weight coefficients of every component of information and communication competencies have been defined by the expert evaluation method: $a=0.2$ (information and communication competencies highlighted by the Federal State Educational Standard); $b=0.3$ (information and communication competencies highlighted by employers); $c=0.25$ (value and motivation component), whereas $a+b+c+d=1$. 15 teachers of Information Technology and subject disciplines, managers of IT departments of the firms where students did placements and students themselves have acted as experts.

This research has defined the formula for the integrated coefficient of the development level of information and communication competencies:

$$K=(aK_1+bK_2+cK_3+dK_4)*100\% \quad (1)$$

where K_1, K_2, K_3, K_4 are development coefficients of the components highlighted by us.

The development coefficient of one of the components is defined by the following ratio:

$$K_1=n_1/n \quad (2)$$

where n_1 is the number of components absorbed by a student; n is the total number of highlighted components.

Development levels of information and communication competencies of future IT professionals have been defined by applying the expert evaluation method: $K < 70\%$ —user level, K between 70% and 90% —technological level, K over 90% —professional level.

It should be noted that a student with a user level means that they only have information and communication competencies specified in the Federal State Educational Standard of Higher Vocational Education. Technological and professional levels include, apart from the standard competencies, the employer's competencies, which, in their turn, are comprised of value and motivation and reflexive and planning units.

In the course of experiments and appraisal survey the experts and students were offered diagnostic tables. The table analysis brings us to the conclusion that most students possess technological level of professional competencies and information and communication competencies. It means that future IT professionals can use information and network technologies in their career, even when the conditions of IT environment are uncertain. The pilot work on the model and technology of qualimetric assessment of professional competencies has enabled us to improve the design and engineering course, to boost the students' motivation and to provide objective and valid presentation of students' and graduates' education results. The effectiveness of results has been proved by statistic processing of pedagogical research results.

4. Discussion

Let us examine the characteristics that are automatically generated by Moodle 2.4 based on the test results and that we believe to be most important in assessment of professional competencies and information and communication competencies.

The scores of a student's knowledge, expertise, skills and competencies in a subject is defined by tests. The scores a student gets after tests are of equal value and can range from 0% to 100% . Let us set the major scores:

P_{ij} - test score. It is defined when a student takes a test in Subject j of Unit i P_{ij} ($0-100\%$) (Units – U_1, U_2, U_3).

Depending on the number of hours allocated for the subject by the Federal State Educational Standard, Unit i is assigned a weight factor for subject j V_{ij} ($0-1.0$)

$$S_{ij} = P_{ij} \cdot V_{ij} \quad (3)$$

is a score for knowledge, expertise, skills and competencies in Subject j of Unit i . S_{ij} ($0-100$).

$$C_i = V_i \cdot \sum_{j=1}^{n_i} S_{ij} \quad (4)$$

is a score for knowledge, expertise, skills and competencies in Unit i . $C_i \in (0-100)$.

$$K = \sum_{i=1}^m C_i \quad (5)$$

is a score for knowledge, expertise, skills and competencies of an engineering university graduate.

This model does not require a threshold value; the weight of each unit and subject is taken into account; quality levels are defined by a competency score enabling us to judge competency or incompetency.

Obviously, while measuring and assessing the quality of a graduate's education, it is impossible to consider a virtually endless diversity of their professionally significant characteristics. Diverse as they may be, it is essential to select only those qualities that are related to satisfying certain social professionally significant needs.

The system of parameters for qualimetric assessment shall be defined based on the compulsory minimum educational programme under the existing Federal State Educational Standard. As each unit is a combination of subjects, it is these subjects that will act as parameters in our model, their content and volume being defined in

the programme.

As the Federal State Educational Standard determines the number of credits/hours for each unit and subject depending on specialism, we turn our attention to the qualimetric assessment of knowledge, expertise, skills and competencies of a graduate specialising in Applied Information Science in Economics.

Having established a structure of each unit, we can get down to calculating the weight of its components, i.e. subjects, based on the Federal State Educational Standard. The curriculum specifies the number of hours for professional unit of subjects U3. Consequently, weights V_{3j} of each competency in U3 are defined as a relation of hours allocated for the subject to total hours of U3. The same procedure is performed for other units. As a result, a hierarchy tree of qualities is drawn, the weights of all qualities are defined. The next step in a qualimetric assessment is to determine individual scores for each student in all the qualities (competencies and their components). It has been decided to define a student's individual scores in competencies and their components by means of Moodle tests, which are used to test students intramurally.

All characteristics are divided into two groups: the first relates to the test as a whole, the second – to each question or group of questions in the test.

The characteristics enabling us to assess the test quality are as follows:

- Average score of testees: arithmetic mean of scores of all students taking the test.

$$T = \frac{1}{S} \sum_{s \in S} T_s \quad (6)$$

where S is students that has taken the test. T_s is a test score of Students.

- Median is a median value of testees' scores T_s . It is this technique that forms the basis of the final table of a test user.

- Standard deviation of test scores: it defines test gradability, i.e. its ability to divide the testees by their level of education. This characteristic is defined by the formula:

$$SD = \sqrt{V(t)} = \sqrt{\frac{1}{S-1} \sum_{s \in S} (T_s - T)^2} \quad (7)$$

- Asymmetry and excess coefficients: measures characterising the difference of distribution structure of testees' scores from normal distribution.

- Test reliability coefficient, aka Cronbach's alpha.

$$CIC = 100 \frac{p}{p-1} \left(1 - (s-1) \frac{\sum_{p \in P} (x_p(s) - \bar{x}_p)}{\sum_{s \in S} (T_s - \bar{T})^2} \right) \quad (8)$$

where p is the test task number out of a group of test tasks P. It is a mean scatter of results of each student when giving answers to all test questions. The less the scatter of the results in relation to the scatter of the total test scores is, the more coherent the test questions are.

- Standard mistake: assesses the factor of luck and defines the bounds of error for the student's test score:

$$SE = \frac{ER}{100} SD \quad (9)$$

If a standard mistake equals to 10% and the student got 60% of the highest score, then his true score will be between 50% and 70%.

The theory of pedagogical measurements states that in a better test the testees' average arithmetic score equals to the median score of the used tasks; asymmetry and excess coefficients do not deviate from the values of a

standard curve of normal distribution of results. It is also good if average arithmetic value, mode and median value coincide. It indicates that the test difficulty level has been accurately adapted to the testees' education level. The concurrent average of scores, asymmetry and excess values also allow us to correctly compare the distribution of results among different tests.

Another group of parameters allows us to assess the quality of concrete test tasks (questions): Index of Ease is percent of students that gave the correct answer to a certain test question. For test task i , it is defined by this formula

$$F_i = \bar{x}_i \quad (10)$$

given a scale of 100 is used, where averaging is done based on all the testees, who did the task; Standard Deviation characterizes the scatter of the testees' scores when answering a particular test question; "Guessed It" Score is a score that a student could get if he/she guessed the answers; Assumed weight is the weight that the lecturer assigned to the test task when forming the test scenario; Effective weight reflects the actual share of a particular question in the students' final score.

The question quality has been statistically assessed by Moodle tools. The essential statistic feature that grades test task abilities and that can be calculated by Moodle is Differentiation Coefficient (DC). The task is considered as having sufficient gradability if the differentiation coefficient is 30% or more.

The DC analysis for the case study revealed that 14% of the test tasks used in the experiment did not satisfy those requirements ($DC < 30\%$). Moreover, this coefficient in one of the tasks is negative (-40.69%). It demonstrates obvious defects of the question. The statistic assessment of test questions database allows you to correct the test.

The most important feature of a test is its reliability. It defines reproducibility and accuracy of results. The reliability coefficient is a correlation coefficient showing the coincidence degree of the results of the same test taken in the same conditions. Test reliability depends on a measurement error. If there is no error, the reliability coefficient is 1. If the measured test score was fully caused by a measurement error, then the test reliability is 0. According to the statistic assessment of the analysed test, the error came to 5.66%, whereas the coefficient of internal consistency was 95.90%.

Considering the information about the test task quality received before, the test was analysed statistically, "incorrect" questions in different parameters being excluded.

References

- Eremina, I., & Sadykova, A. (2013). Theoretical foundations and principles of building education information environment of federal university training IT professionals and its implementation. *Educational Technologies and Society*, 3. Retrieved from http://www.ifets.ieee.org/russian/periodical/V_163_2013EE.html.
- Eremina, I., Gareeva, G., & Mingazova, G. (2013). Information educational environment of the Federal University as one of the leading directions of the improvement of the modern system of professional training of future IT-professionals. *Eastern-European Scientific Journal*, 3. Retrieved from <http://www.auris-verlag.de>
- Grice, H. P. (1975). Logic and conversation. In P. Cole, & J. L. Morgan (Eds.), *Syntax and semantics* (Vol. 3, pp. 41-58). N.Y.: Academic Press.
- Hawkes, L. W., Derry, S. J., & Rundensteiner, E. A. (1990). Individualized Tutoring Using an Intelligent Fuzzy Temporal Relational Database. *International Journal of Man-Machine Studies*, 33(4), 409-429. [http://dx.doi.org/10.1016/S0020-7373\(05\)80040-9](http://dx.doi.org/10.1016/S0020-7373(05)80040-9)
- Polad, E., Bukharkina, M., Moiseyeva, M., & Petrov, A. (2009). *New pedagogic and information technologies in education*. Akademiya Publishing Centre.
- Starichenko, B. (2004). *Computerised Treatment and Presentation of Data of Pedagogical Research*. Ekaterinburg: State Teacher Training University.

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/>).