KONSTANTIN PRESLAVSKY UNIVERSITY OF SHUMEN

The Faculty of Mathematics and Computer Science

Department Computer Systems And Technologies

BORYANA UZUNOVA-DIMITROVA
HRISTO PARASKEVOV
STANIMIR ZHELEZOV

OBJECTIVE ASSESSMENT DURING A PANDEMIC

Training over the past year has posed serious challenges posed by the global COVID-19 pandemic, necessitating comprehensive distance learning and subsequent distance assessment. In practice, the classical methods of training and assessment have been replaced by the use of information and communication technologies. The problem of objective assessment is informal and is solved in conditions of uncertainty of various types, determined by criteria and language rules for decision-making.

The purpose of the present study is to analyze the results obtained, as the learning material is transmitted entirely online, in the case of videoconferencing via Google Meet. The experiment was conducted with 52 students studying the discipline "Web technologies". The verification of the acquired knowledge, skills and competencies is carried out by means of an electronic test and a developed course project.

The automation of control procedures implies formalization and adequate modeling of the teacher's thinking in the decision-making process for the formation of the assessment.

For this purpose, we use criteria representing the linguistic variables, through which one can more easily evaluate, as each criterion takes a value from a predetermined term-set {very bad, bad, not very good, good, very good, excellent, perfect}.

For formalization of the task and interpretation of the results of the criterion-oriented testing and evaluation of

the quality of the course project the following solution is proposed.

$S=\{s_1, s_2, ..., s_m\}$ - finite, discrete set of persons subject to assessment;

D - finite, discrete set of diagnoses, final grades {2, 3, 4, 5, 6};

$T=\{t_1, t_2, ..., t_n\}$- finite, discrete set of expertly set test with test questions / criteria t_n, n=1, 2, ..., k;

$A = ||a_{ij}||$, i=1,2,...m, j=1,2,...n - matrix containing the results of the evaluation, where $a_{ij} \in L$ is the evaluation of the s_i-th work on the t_j-th test question / criterion;

L = {0, 1} - discrete scale of values of the scores of each individual test question or criterion.

Therefore, the task of interpreting the results could be formulated as follows: "**For each evaluated person $s_i \in S$ to determine a final grade $d \in D$,**

based on the results obtained in A from

the evaluations of the test questions / criteria $t_j \in T$ set on scale L."

Formally, this means to find an injective image:

$$\Omega: S \to A \to D$$

of quantitatively measured opinion of the quality of the written works A in the set of diagnoses D. The image $S \to A$ is obtained as a result of evaluation of works $s_i \in S$ on the basis of given criteria $c_j \in C$ by the lecturer. In order to obtain $A \to D$ we apply the Rasch model for dichotomous data.

To objectify the assessment, the one-parameter Rush model is used, which is one of the main methods of formalization in probabilistic modeling for measuring student achievement (IRT). It establishes a correspondence between the observed results of the test performance and two sets related to the difficulty of the test / performance of the relevant criteria (T) and the preparation of the learners (S).

To apply the Rush model, we assume that the learner's preparation and the complexity of the test / course project are assessment parameters that allow for objective assessment.

To paraphrase Rush's model, we can assume that the probability P for a student with experience S to answer a test question / create a high-quality project with difficulty T is determined by the following formula:

$$\mathrm{P}(\mathrm{S},\mathrm{T}) = \frac{S}{\mathrm{S} + T}$$

Rasch's model is applied with the dichotomous scale {yes, no} ≡ {0, 1}, which is a bit informative. We choose the scale L = {bad, good, excellent} ≡ {0, 0.5, 1}, which is applicable with minor modifications of the Rasch model. This scale is convenient for teachers and gives a greater opportunity to form an unambiguous assessment compared to the scales in which the number of terms k> 3.

In fig. 1. the characteristic curves of the students in relation to their test score are shown, and in fig. 2. - according to the course project developed by them. It is observed that the curves differ only in their location, ie. the curves are parallel and not intersecting as is characteristic of the characteristic curve of the one-parameter Rush model in IRT theory.

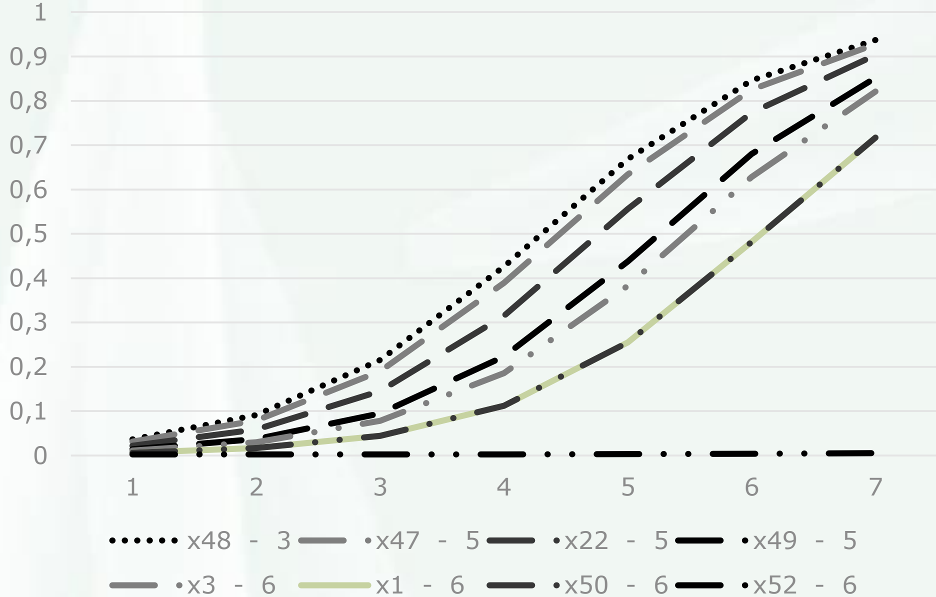

Fig. 1. Achieved test results against the Rush assessment

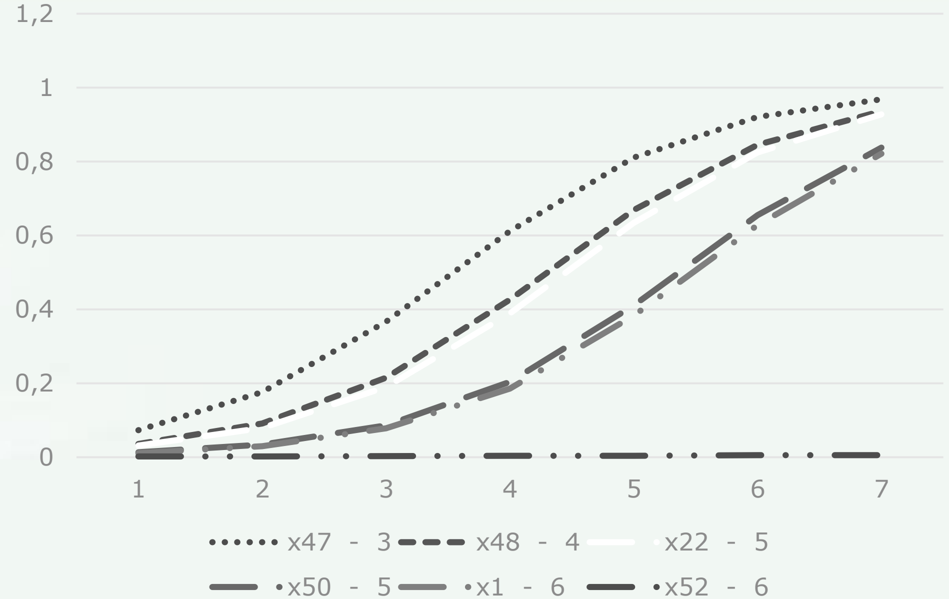

Fig. 2. Achieved results of a course project compared to the Rush assessment

Analyzing the results, we can draw the following conclusions:

- The proposed model implements a multicriteria approach for measuring the quality of the test / course project of students.
- Test questions / criteria must allow for an objective assessment. Therefore, they must be accurate, clear, unambiguous, without the use of logical connections (and, or, not).

- The described algorithm allows obtaining an approximate assessment of the test / course project. Assessment according to the Rush model is favorable to knowledgeable students, answering test questions independently or developing their own course projects, their characteristic curve is completely concave (x47, x48) and therefore unfavorable to ignorant, whose characteristic curve is convex, and students with standard characteristic curve for the Rush model or parallel to the abscissa (x1, x22, x50, x52) correspond to well-prepared students.
- The results of the thus assessed tests / course projects of the students shown in Figures 1 and 2 show that the assessment method used is reliable and objective.

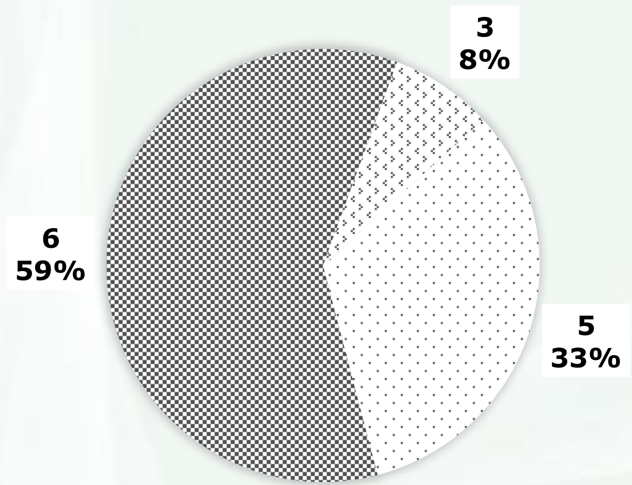

Fig. 3. Ratio of test results achieved and number of students

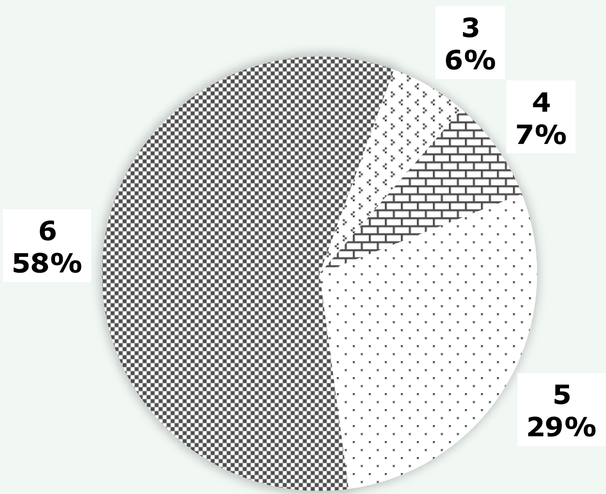

Fig. 4. Ratio of achieved results of the course project and number of students

Figures 3 and 4 show the ratio of the achieved results of the test and the presented course project and the number of students, compared to the proposed modified Rush method. It is observed that the obtained results overlap or approach each other, which in turn shows that the assessments of the learners are objective to their acquired knowledge. The evaluation of the test is comparable and realistic with the evaluation of the course project.

The expectations for overlapping the grades from the two types of tests are justified, which is shown in Figure 5 where the achieved results of the learners are presented. It is clear from the figure that there are no drastic differences between the two types of evaluation.

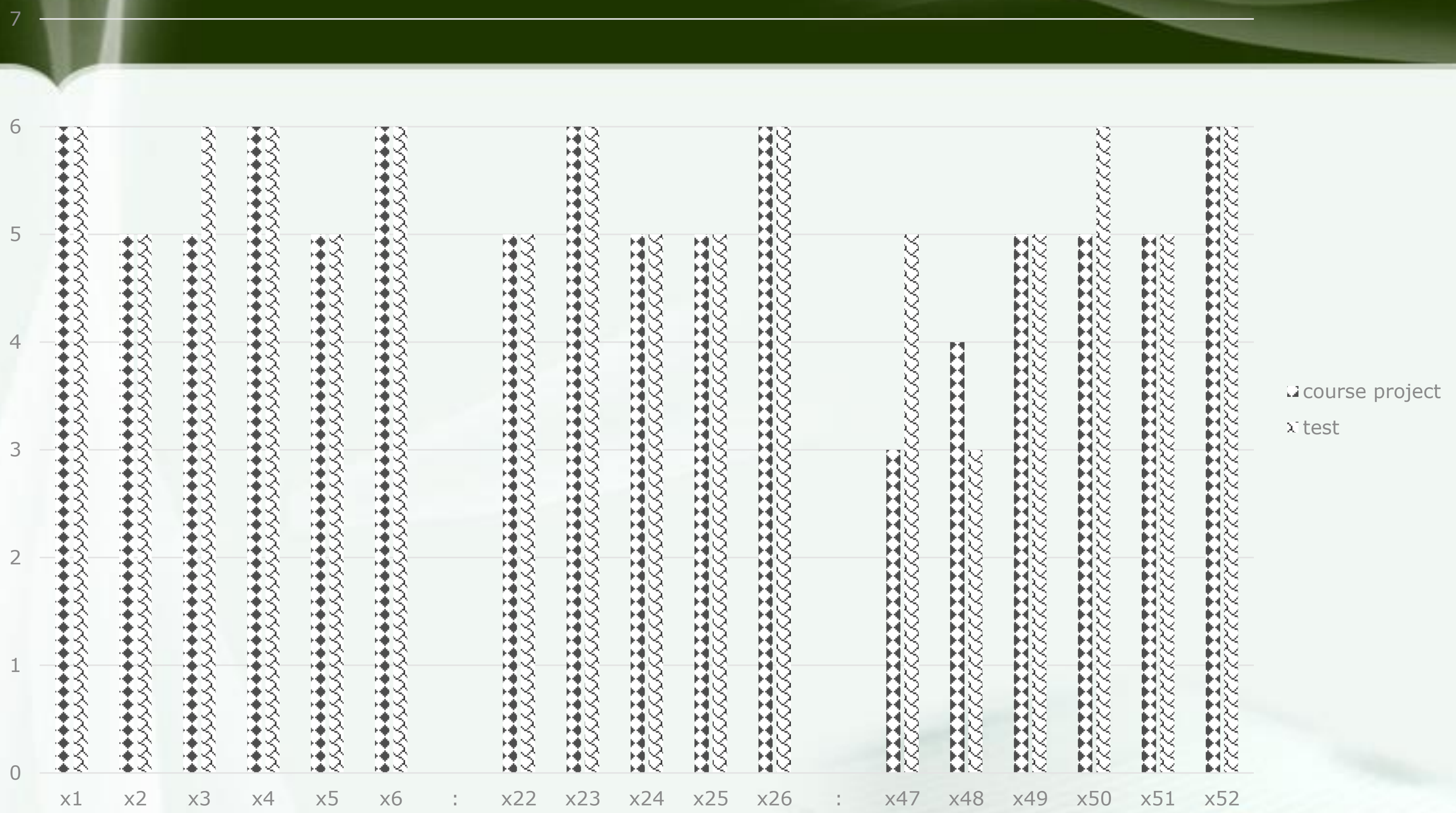

Fig. 5. Achieved results when testing a practical task and through a test

The proposed method for knowledge assessment could be implemented in software as part of a developed "Web-based intelligent knowledge management system for students", which will facilitate teachers for objective assessment electronically.

THANK YOU FOR YOUR ATTENTION!