PERIÓDICO TCHÊ QUÍMICA
ARTIGO ORIGINAL

AVALIAÇÃO BASEADA EM CRITÉRIOS COMO MODO DE FORMAÇÃO DA ALFABETIZAÇÃO FUNCIONAL DOS ESTUDANTES EM CIÊNCIA DA COMPUTAÇÃO

CRITERIA-BASED ASSESSMENT AS THE WAY OF FORMING STUDENTS' FUNCTIONAL LITERACY IN COMPUTER SCIENCE

КРИТЕРИАЛЬНОЕ ОЦЕНИВАНИЕ КАК СПОСОБ ФОРМИРОВАНИЯ ФУНКЦИОНАЛЬНОЙ ГРАМОТНОСТИ УЧАЩИХСЯ ПО ИНФОРМАТИКЕ

AVDARSOL, Sailaugul[1*]; RAKHIMZHANOVA, Lyazzat B.[2]; BOSTANOV, Bektas G.[3]; SAGIMBAEVA, Ainur Ye.[4]; KHAKIMOVA, Tiyshtik[5];

[1,3,4] Abai Kazakh National Pedagogical University, Institute of Mathematics, Physics and Informatics, Department of Informatics and Informatization of Education – Republic of Kazakhstan

[2,5] Al-Farabi Kazakh National University, Department of Computer Science – Republic of Kazakhstan

** Corresponding author*
e-mail: sailau-abai@mail.ru

Received 14 February 2020; received in revised form 03 April 2020; accepted 24 April 2020

RESUMO

Durante muito tempo, a principal abordagem da avaliação foi a abordagem normativa, quando as realizações individuais dos estudantes foram comparadas com uma determinada norma (os resultados da maioria dos estudantes). Recentemente, a pesquisa pedagógica doméstica tem desenvolvido uma abordagem baseada em critérios para avaliar o desempenho acadêmico, de acordo com que as realizações dos estudantes são comparadas com o escopo de conhecimento que deve ser adquirido em um estágio específico do treinamento. Este estudo teve como objetivo determinar o papel da avaliação baseada em critérios na formação da alfabetização funcional dos estudantes em ciência da computação e construir um modelo de avaliação baseada em critérios no desenvolvimento da alfabetização funcional além de demonstrar a eficácia dos métodos de formação dos alunos. 'alfabetização funcional em ciência da computação. Os principais métodos de pesquisa foram a metodologia de avaliação baseada em critérios desenvolvida pelos autores e a metodologia de avaliação formativa. Foram analisados alguns elementos da metodologia de avaliação formativa. Para o desenvolvimento adicional da metodologia para a formação da alfabetização funcional dos estudantes em ciência da computação, foi construído um modelo de avaliação baseada em critérios. A introdução da avaliação baseada em critérios permitirá mudar para uma avaliação formativa destinada ao desenvolvimento da competência do estudante. A avaliação, composta por critérios que o estudante entende, o estimula e torna o processo de aprendizado lógico. Com base nos experimentos práticos e na avaliação proposta com base em critérios, foi comprovada a eficácia de métodos para a formação da alfabetização funcional dos estudantes em ciência da computação.

Palavras-chave: *avaliação crítica, educação, resultados, avaliação, ciência da computação.*

ABSTRACT

For a long time, the primary approach to assessment was the normative approach when the individual achievements of students were compared with a particular norm (the results of most students). Recently, domestic pedagogical research has been developing a criteria-based approach to assessing academic achievement when students' achievements are compared with the amount of knowledge that needs to be acquired at a particular stage of training. This study aimed to determine the role of criteria-based assessment in the formation of students' functional literacy in computer science and to build a criteria-based assessment model in the development of functional literacy beyond to demonstrate the effectiveness of the methods of formation of students' functional literacy in computer science. The leading research methods were the criteria-based assessment methodology developed by the authors and the method of formative assessment. Some elements of the methodology of forming evaluation were considered. For further development of the methods for the formation of students' functional literacy in computer science, a criteria-based assessment model has been built. The introduction of criteria-based assessment will allow to switch to a formative evaluation aimed at developing student competence. The evaluation, consisting of criteria that a student understands, stimulates him and makes the learning process meaningful. Based on practical experiments and the proposed criteria-based assessment, the effectiveness of

DOI: 10.52571/PTQ.v17.n35.2020.04_AVDARSOL_pgs_41_54.pdf

methods for the formation of students' functional literacy in computer science has been proved.

Keywords: *critical assessment, education, results, assessment, computer science*.

АННОТАЦИЯ

На протяжении долгого времени основным подходом к оцениванию был нормативный подход, когда индивидуальные достижения учащихся сравнивались с определенной нормой (результатами большинства обучающихся). В последнее время в отечественных педагогических исследованиях разрабатывается критериальный подход к оценке успеваемости, когда достижения учащихся сравниваются с объемом знаний, который подлежит усвоению на определенном этапе обучения. Целью данного исследования было определить роль критериального оценивания при формировании функциональной грамотности учащихся по информатике и построить модель критериального оценивания для развития функциональной грамотности, чтобы показать эффективность методов формирования функциональной грамотности учащихся по информатике. Ведущими методами исследования стали разработанная авторами методика критериального оценивания и методика формирующего оценивания. Рассмотрены некоторые элементы методики формирующего оценивания. Для дальнейшей разработки методики формирования функциональной грамотности учащихся по информатике, построена модель критериального оценивания. Внедрение критериального оценивания позволит перейти к формативному оцениванию, направленному на развитие компетентности учащегося. Оценка, состоящая из критериев, понятных ученику, стимулирует его и делает процесс обучения смысловым. На основании практических экспериментов и предложенного критериального оценивания доказана эффективность методов формирования функциональной грамотности учащихся по информатике.

Ключевые слова: *критическое оценивание, образование, результаты, оценивание, информатика.*

1. INTRODUCTION:

For a long time, the main approach to assessment was the normative approach, when the individual achievements of students were compared with a particular norm (the results of most students) (Yang *et al.*, 2017; Yang *et al.*, 2019). Recently, domestic pedagogical research (Krasnoborova, 2009; Sokolova, 2017; Shalashova *et al.*, 2018; Totikova *et al.*, 2019) has been developing a criteria-based approach to assessing academic achievement, when students' achievements are compared with the amount of knowledge that needs to be acquired at a certain stage of training since there is a need for assessment, the results of which would allow to determine what educational goals a student has achieved. This led to the advancement of such an aspect of assessing as comparing an individual result with predetermined criteria.

Evaluation systems in different countries show that a five-point grading system with vague indicators is traditional and does not satisfy modern educational needs (Müller *et al.*, 2017; Cencelj *et al.*, 2019; Guo and Yan, 2019; Kim, 2019). A multi-point system is conventional, but different countries practice different points; the most optimal, in authors' opinion, is the letter designation of achievements, which means a certain number of points (Haber and Mitchell, 2017; Lookadoo *et al.*, 2017; Buchholtz *et al.*, 2018; Sagimbaeva *et al.*, 2019).

Kazakhstani school education is being developed in accordance with the State Program for the Development of Education of the Republic of Kazakhstan for 2011-2020. The updated educational program is aimed at developing a wide range of functional literacy of students, following international practice, enables students to develop functional literacy and critical thinking. In the program of Kazakhstan on the development of the functional literacy of students for 2012-2016, the results of the development of functional literacy of students and the application of acquired knowledge in practice are indicated as the main competencies.

Currently, Nazarbayev Intellectual Schools JSC, together with the Cambridge University Board of Examiners is developing experimental learning outcomes for each section and topic of the subject in innovative integrated training programs, and ways to assess their achievement. In a primary school (7–10th grades), students learn the same program material over a different period, depending on their abilities, and by the end of the 10th grade, they must achieve the same learning goals. This approach to the organization of the educational process, as it is seen, gives a high result (Sagimbaeva *et al.*, 2019) of training, and in the future, leads to a successful adult life.

In the curriculum for assessing the performance of students of other secondary schools, in educational materials, practical experience systematically affecting the formation of educational and cognitive competence of students, the established criteria for assessing the achievements of students, who have interdisciplinary sciences, are not defined in advance. Criteria-based assessment in the educational process is not fully understood. Existing assessments and monitoring do not allow to determine whether achievements in the field of education are fully used since the subjective assessment tools used (tests, theoretical questions, control issues, etc. are not criteria for determining the level of knowledge acquisition) (Broadbent *et al.*, 2018; Sagimbaeva *et al.*, 2019)

The establishment of various approaches to assessing competence, personal orientation, the development of a general pedagogical concept based on the development of relationships – all this determines the current trends in the development of the assessment system, and some criteria for assessing the level of required competence are based on a comparison of students' achievements, methodological problems of criteria-based assessment (Carter and Bathmaker, 2017; Sanchez *et al.*, 2017; Alt, 2018; Mohamadi, 2018; Astriawati and Djukri, 2019; Basera, 2019; Palmiero and Cecconi, 2019; Mirmotahari *et al.*, 2019).

Thus, based on the theoretical and practical foundations of world and domestic experience, it can be concluded that as part of the educational process, assessment and control and evaluation activities as a whole act as an independent element of the content of education, requiring development (Azmy and Mokhtar, 2017; Marinkovich *et al.*, 2018; Mohamadi Zenouzagh, 2019; Pastore *et al.*, 2019; Silseth and Gilje, 2019). At the same time, the assessment system allows to receive integral and differentiated information about the educational process, track individual student progress in achieving the planned results, provide feedback for teachers, students, and parents, track the effectiveness of the educational program (Luckin *et al.*, 2017; Meek *et al.*, 2017; Houston and Thompson, 2017; Black and Wiliam, 2018; Tienson-Tseng, 2019).

This study aimed to determine the role of criteria-based assessment in the formation of students' functional literacy in computer science and to build a criteria-based assessment model in the development of functional literacy beyond to demonstrate the effectiveness of the methods of formation of students' functional literacy in computer science.

2. LITERATURE REVIEW:

A.A. Krasnoborova (2009), in her article, proposes the author's definition of pedagogical technology for criteria-based assessment. The authors call the pedagogical technology of criteria-based evaluation of students in the framework of educational and cognitive activity a process-effective meta-technology that provides a system of interconnected control and assessment of actions of all participants in the educational process to achieve the goals and objectives of the training.

In the study of E.A Sokolova. (2017) a methodology is developed in which criteria-based internal assessment is included in all stages of the activity-oriented design of teaching the topic of the school geometry course developed in this study. Studies suggest different approaches to assessing results.

The study by M.M. McGill (McGill and Decker, 2019) introduces a Resource Centre (csedresearch.org) designed to provide resources for researchers in primary and secondary computer education. The center has a primary function to provide a centralized location for assessment tools, many of which focus on computing.

Y. Ohashi (Ohashi and Yamachi, 2017) used participant observation and group interviews as research methods, as well as a case code matrix and a qualitative data analysis method. The authors developed assessment criteria for an information volunteer for a training course launched in 1997, in which students participated in educational activities based on information from nearby schools. The study generated 30 cases, which were organized in the following three categories: problems, support, and what they learned from this activity. Each of these categories was further divided into the following subcategories: communication, office work, and training. It was discovered that students could not solve all the problems. Also, students demonstrated different levels of performance, which, according to their conclusion, may be associated with differences in students' academic abilities, as well as with the diversity of activities from school to school.

When determining functional online literacy in the study by K. Dolenc (Dolenc *et al.*, 2015), it is demonstrated that there are differences in reading comprehension when reading offline and online, when using electronic school material in their educational process. The study involved 78 students of the 8th grade of primary school, studying the course “Technology and Science.” The authors used individual and adaptive intelligent learning systems (ITS) and, using the assessment of the results, showed that for this form of ITS there is still enough room for optimization, which is a constant method of improvement and updating in such systems.

Priscilla Haring (Haring *et al.*, 2018), examining information on the relationship between game design and player’s cognitive processes, explores the use of Bloom's taxonomy in describing a psychotherapeutic game in terms of knowledge and cognitive processing.

The study by Valerie J. Shute (Shute *et al.*, 2011) examines the research literature on the relationship between parental involvement (PI) and academic performance, with a particular emphasis on high school (middle and high school). The results first show how individual PI variables relate to academic achievement, and then move on to a more sophisticated analysis of several variables of the general design described in the literature.

Yuan Wang (Wang *et al.*, 2018), in his studies, showed that assessing the skills of information literacy students is extremely important. This article explores an improved compression criterion, which is based on the entropy of a population in an objective space and the maximum distance in a decision-making space and is used to make a decision about starting a local search. The authors argue that the modified model-based population reinitialisation strategy is designed to enhance the ability of HDEMR (hybrid differential evolution with model-based reinitialisation) to search globally for solving complex problems.

Panagiotis Psomos (Psomos and Kordaki, 2011) presents a new model for assessing the pedagogical validity of educational digital storytelling environments (EDSE). This model is based on modern social and constructivist views on learning and consists of sixteen dimensions. The model has the shape of a star called the “Digital Storytelling Pedagogical evaluation star,” consisting of sixteen peaks as the number of aspects of the assessment above model. The proposed model can help researchers and designers in the pedagogical analysis of the existing EDSE and in making appropriate decisions for the design of the future EDSE. Besides, this model can help teachers choose the right EDSE so that they can complete specific pedagogical tasks in their classrooms.

Rosário R. (Rosário *et al.*, 2011) analyses the role played by some homework variables in student achievement (proximal and distal), as well as their mediating role in using self-regulatory learning strategies and perceived self-efficacy in the subject area. The focus is on English as a Foreign Language (EFL), and a sample of 591 Portuguese fifth and sixth-grade students is used. The data they obtained confirm the indirect effect of homework on school achievement with the help of the mentioned cognitive and motivational variables (the use of self-regulatory learning strategies and self-efficacy).

Inga Glogger (Glogger *et al.*, 2013) shows in her study teacher preparation for evaluating important components of self-regulatory learning, such as learning strategies, which is an essential aspect of integrating self-regulatory knowledge in the school. The authors argue that training journals can be used to evaluate learning strategies per models of the cyclic process of self-regulatory learning and provide precious formative feedback. A computer-based learning environment (CBLE) has been developed that educates teachers in evaluating learning strategies using instructional journals.

The goal of Manuela Leidinger (Leidinger and Perels, 2012), based on Zimmermann’s theory of self-regulation, was to promote a robust learning environment to support self-regulatory learning using teaching materials. Learning materials have been developed that focus on the specific (meta) cognitive and motivational components of self-regulatory learning and are divided into six blocks, with which students from the experimental group worked weekly. In total, 135 fourth-graders took part in the study.

The article by A. Vernon (2011) reviews a model developed by the authors to evaluate non-traditional teaching methods, such as group learning. This model was applied to group training courses, which included innovative enhancements such as group exams and group-based role-playing games. The model required a balanced presentation among the set of learning criteria, grouped into six sets of target criteria; all of them are taken from the literature on group learning. The authors developed a test tool based on this model and presented it to 85 students at the end of three business strategy courses.

Many topical questions were asked by teachers in a study by Vreda Pieterse (Pieterse and du Toit, 2009) in a closed electronic survey to identify possible causes for misrepresentation of questions in assessing knowledge, due to the students' inability to sufficiently understand their essence. In this study, they tried to extrapolate aspects that could lead to computer science students misunderstanding instructions and scientific research. As a result, they would not be able to solve the very questions that the lecturer wants to evaluate. The authors sought to compile a set of guidelines to address issues related to misunderstanding of the issues raised.

Functional literacy consists of language, legal, environmental, computer, information, and activity literacy. During the current continuous development of the information society, everyone has the opportunity to master information technology and use it according to their own needs, that is, forms their own computer literacy, which is one of the components of students' functional literacy, and it is also known that computer science plays a huge role in the formation of information and technological competencies, which are an integral part of the competencies indicated above.

The expected results of students in computer science provide an opportunity to objectively evaluate their educational achievements and determine their development paths, considering the abilities of each student, also help stimulate students' skills and expect improvement in the quality of the educational process.

In reflecting the assessment of computer science for K-8 teachers, Hannah E. Chipman (Chipman *et al.*, 2019) concluded that one area that has not been fully explored is how grades can affect student confidence and attitudes to CS. It was shown that the threat of stereotype and modality of a test affect the performance of the tested CS.

As a result of studying the functional literacy of students in the work of Anatoly Veryaev (Veryaev *et al.*, 2012), it was shown that functional literacy reflects practical ideas about adaptation to social conditions, and cannot be considered as one of the leading educational goals. It is concluded that the natural-scientific direction of training, in comparison with the humanities, has a more favorable effect on the test results in the framework of tasks proposed by international organizations. As a result, the identification of computer and information literacy levels allowed to test the developed test tasks focused on the Federal State Educational Standard to adjust lectures and practical classes for first-year students at the Computer Science course.

The monograph by H.S. Bhola (1979) considers assessment as a link in the "ideology for technology" chain, assessment models and approaches, an assessment model for a specific situation, providing assessment in functional literacy programs, classic change measurement protocols, some newly discovered change measurement protocols, implementation data for management and evaluation, particular assessment tasks for employees of functional literacy, data verification, and analysis, design of assessment and interpretation of evaluation results for decision making.

3. MATERIALS AND METHODS:

Criteria-based self-assessment: at the beginning of work with this technique, a teacher needs to determine the criteria for assessment with students. A teacher should show by example the levels of achievement of these criteria and how the assessment process will take place. It is also useful to practice evaluating works. Examples of criteria-based self-assessment tools are systems of conditional signals and designations of the level of understanding or mastering by a student of a particular knowledge or skill (for example, "color tracks", "traffic lights", "rulers", "cards +/-").

For criteria-based mutual assessment, maps or sheets of assessment are developed with criteria and levels of achievement or forms of feedback on a work. Weekly reports are sheets that students fill out at the end of the week, containing three questions: What have I learned this week? What questions remained unclear for me? What questions would I ask students if I were a teacher to check if they understood the material?

The mind map (intellect map, concept map) consists of the names of concepts placed in frames; they are connected by lines fixing the relations of these concepts from general to particular. A map is considered from top to bottom, a teacher will be able to determine how well students see the overall picture of an entire course or a single topic.

The introduction of criteria-based assessment will allow switching to a formative evaluation aimed at developing student competence. The assessment, consisting of criteria that a student understands, stimulates him, and makes the learning process meaningful. The criterion approach to assessing students consists in comparing the student's achievements with well-defined, collectively developed criteria well known to all participants in the process. Criteria of assessment are developed for each subject. With proper preparation of the criteria scale, the student can independently assess the quality of their work, which stimulates the achievement of a higher educational result and the formation of educational independence. The use of criteria-based assessment in the educational system makes it possible to identify and improve the method for assessing student performance with the objective goals of an individual subject, as well as using specific parameters (criteria) that allow students to compete in high school.

4. RESULTS AND DISCUSSION:

Criteria assessment consists of formative and summative assessments. The training manual of the Ministry of Education and Science of the Republic of Kazakhstan (Methodological and educational..., 2017) indicates that formative assessment is an assessment of the current work in the classroom, which is mandatory for this stage of training and the cognitive process. This is the student’s current performance indicator, providing feedback between a student and teacher. It helps to identify student difficulties and increase the ability to achieve good results. A student should be able to fill current gaps in learning, considering the recommendations of a teacher, with the advice of other students, completing the remaining or additional tasks.

Summative assessment is a cumulative form of assessment that is carried out at the end of a specific training period (quarter, trimester, academic year), and also after mastering the section (Methodological and educational ..., 2017). Following the curriculum, the criteria-based assessment is conducted continuously by a teacher, provides feedback between a teacher and students, and allows timely correction of learning processes, not only scores, grades. A total assessment of sections/general topics is carried out by a teacher 2-3 times a quarter. The final quarterly assessment is carried out by a teacher at the end of each quarter. Summative assessment to determine the level of knowledge is carried out at the end of the primary, principal, senior grade. The funded system records only the achievements of students; therefore, actively encourages students to self-study and cognitive activity, to the full mastery of the curriculum. Standards, methods, and means of assessment vary depending on the type of assessment and subject specificity. The authors present the developed criteria-based evaluation model in the form of a scheme (Figure 1). Consider the techniques of formative assessment (Figure 2). Each of them is based on different criteria.

Computer science training mainly involves the development of user skills, in particular computer skills and new technologies. Therefore, when studying some topics, practical work on the computer is conducted. When developing and testing experimental integrated educational programs in subjects, the authors partially began to use the approach used by Nazarbayev Intellectual Schools in conjunction with the Examination Council of the University of Cambridge.

The authors want to focus on the practice of preparing assignments and carrying out SAS (summative assessment of a section) and SAQ (summative assessment of a quarter) when teaching computer science. When developing tasks of SAS and SAQ with the use of criteria-based assessment at computer science lessons and meeting the requirements of the modern system of updated content in training, it is necessary to strive to develop students' thinking, to provide an opportunity to analyse the proposed situations, thereby students learn to analyse and evaluate their critical thinking. When developing test tasks in the SAQ, it is necessary to adhere to the test specification (ST), where the learning objectives and tested skills are predefined.

For example, consider the topic “Measuring information and a task” in the “Computer” section (Table 1). It is possible to accept tasks in the form of tests (Table 2).

For an objective and evidence-based verification of the validity of the pedagogical hypothesis, a pedagogical experiment was conducted using the following methodology. As a result of measuring the same indicator using the same measurement procedure, the following data are obtained: $x = (x_1, x_2, \ldots, x_n)$ – sample for the experimental group and $y = (y_1, y_2, \ldots, y_m)$ – for the control group, where x_i – the sample element – is the value of the studied indicator (feature1) for the i-th member of the experimental group, i = 1, 2, ..., n, and y_j is the value of the studied indicator for the y_j -th member of the control group, j = 1 , 2, ..., m. The number of sample elements is called its size — for example, the size of sample x is N, and the size of sample y is M (Novikov, 2004).

Measurements were made on the *ratio scale* (time, number, etc.), $\{x_i\}$ and $\{y_j\}$ – positive, including natural numbers, for which all arithmetic operations make sense. The measurement consists in determining the level of knowledge by conducting a test that includes 20 tasks. It is assumed that the characteristic of a student (sign) is the number of correctly solved problems.

The authors use an *ordinal scale* (rank scale) with L gradations; it is assumed that $\{x_i\}$ and $\{y_j\}$ are natural numbers that take one of L values. A set of values (points) is a set of numbers from one to L. A characteristic of a group will be the number of its members who have gained a given score. That is, for the experimental group, the point vector is $n = (n_1, n_2, \ldots, n_L)$, where n_k is the number of members of the experimental group who received the k-th point, k = 1, 2, ..., L. For the control group, the point vector is $m = (m_1, m_2, \ldots, m_L)$, where m_k is the number of control group members who received the k-th point, k = 1, 2, ..., L. Obviously, $n_1 + n_2 + \ldots + n_L = N$, $m_1 + m_2 + \ldots + m_L = M$.

Here are the formulas for calculating the main indicators. The arithmetic mean x of the sample $\{x_i\}_i$ = 1 ... N (sample mean) is calculated as follows (Equation 1) and the sample dispersion D_x (Equation 2). For the data measured in the ratio scale, to test the hypothesis that the characteristics of the two groups coincide, it is advisable to use the Cramer-Welch criterion. The Cramer-Welch criterion is designed to test the hypothesis of equality of means (strictly speaking, mathematical expectations) of two samples.

The empirical value of this criterion is calculated on the basis of information about the volumes N and M of samples x and y, sample means x and y, and sample dispersions Dx and Dy of the compared samples (these values can be calculated manually using formulas (1)-(2) or using the tool "Descriptive statistics" in the computer program Microsoft Excel according to the following formula (Equation 3):

Descriptive statistics, firstly, allows presenting the results of a pedagogical experiment in a compact and informative form, which makes it possible to conduct a qualitative analysis of the objects studied. Secondly, several indicators of descriptive statistics are used in quantitative analysis (when applying statistical criteria). The algorithm for determining the accuracy of coincidences and differences in the characteristics of the compared samples for the experimental data measured in the ratio scale using the Cramer-Welch criterion is as follows:

1. To calculate for the analyzed samples the T_{emp} – the practical value of the Cramer-Welch criterion according to formula (3).

2. To compare this value with the critical value T0.05 = 1.96: if Temp ≤ 1.96, then conclude: "the characteristics of the compared samples coincide at a significance level of 0.05"; if T_{emp}> 1.96, then conclude "the reliability of differences in the characteristics of the compared samples is 95%".

Experimental training was conducted in several schools in Almaty. Students were divided into experimental and control groups. The purpose of the pedagogical experiment was to verify the effectiveness of the developed system of tasks and tests and the authors' proposed methodology for the formation of students' functional literacy in computer science.

In the experimental and control groups, the SAQ was carried out according to the authors' tasks and tests. The experimental group consisted of N = 15 people, and the control group consisted of M = 20 people. The results of measurements of the level of knowledge in the control and experimental groups before and after the experiment are shown in Table 3.

The experimental results are also obtained in the ordinal scale. For N = 15, M = 20, there are three levels of knowledge L = 3: low (the number of solved problems is less than or equal to 10), medium (the number of solved problems is strictly bigger than 10, but less than or equal to 15) and high (the number of solved problems is strictly bigger than 15). The authors form Table 4, in which the upper limits of the ranges are indicated (Figure 3).

For the measurement results in the ratio scale (Table 4), the descriptive statistics indicators were divided into several groups: position indicators describing the position of the experimental data on the numerical axis; scatter indicators describing the degree of spread of data relative to its center (average value). These include sample dispersion, the difference between the minimum and maximum elements (range, sampling interval), etc.; asymmetry indicators: the position of the median relative to the average, etc.

These indicators were used for visual representation and primary ("visual") analysis of the results of measurements of the characteristics of the experimental and control groups. Descriptive statistics for the first column of Table 3 (the number of correctly solved problems in the control group before the start of the experiment) are shown in Table 5.

The authors review the algorithm for determining the accuracy of coincidences and differences in the characteristics of the compared samples for the data from Table 1, measured in the ratio scale, using the Cramer-Welch criterion (Table 6). To do this, first, it is needed to compare the number of correctly solved problems in the control and experimental groups before the experiment. The authors calculate by the formula (3) the value of $T_{emp} = 0.84 \leq 1.96$. Therefore, the hypothesis that the characteristics of the control and experimental groups coincide before the start of the experiment is accepted at a significance level of 0.05.

Now compare the characteristics of the control and experimental groups after the end of the experiment. The authors calculate by the formula (3) the value of $T_{emp} = 2.04 > 1.96$. Therefore, the reliability of differences in the characteristics of the control and experimental groups after the end of the experiment is 95%.

So, the initial (before the start of the experiment) states of the experimental and control groups coincide, and the final ones (after the end of the experiment) are different. Therefore, it can be concluded that the effect of changes is due precisely to the application of the experimental teaching methodology.

5. CONCLUSIONS:

The study of criteria-based assessment as a way of forming students' functional literacy in computer science and the results of experimental work have led to many conclusions:

1. The necessity to ensure the mastery of the primary content of the subjects studied and the formation of functional literacy among schoolchildren in seven key competencies require the development of a new system for assessment by criteria.

2. The basic content of education involves objective results aimed at enhancing the functional, including practical, focus of training. Despite the different number of points in the assessment and the divergence of opinions on this issue in different countries, in all of these systems the unifying core is the criterion of assessment and the differentiation of the levels of assimilation of educational material of students.

3. The process of developing functional literacy among schoolchildren determines the introduction of a new assessment system that takes into account the effectiveness of all types of educational activities, the procedural side of learning material, and the manifestation of individual and personal qualities of students.

4. When studying in the elementary school (7–10th grades), the same program material for a different period, depending on individual abilities, by the end of the 10th grade, students should achieve the same learning goals. This approach to the organizing the educational process gives a high learning outcome, and in the future, leads to a successful adult life.

5. The assessment system allows to receive integrated and differentiated information about the educational process, track individual student progress in achieving the planned results, provide feedback for teachers, students, and parents, track the effectiveness of the educational program.

As a result of the study, the following results were obtained:

1. For further development of the methodology for the formation of students' functional literacy in computer science, a criteria-based assessment model has been built.

2. Based on practical experiments and the proposed criteria-based assessment, the effectiveness of methods for the formation of students' functional literacy in computer science has been proved.

6. REFERENCES:

1. Alt, D. *Learning Environments Research*, **2018**, *21*(3), 387-406.
2. Astriawati, F., Djukri. *Journal of Physics: Conference Series*, **2019**, *1397*(1), article number 012049.
3. Azmy, A., Mokhtar, H. *International Journal of Engineering Education*, **2017**, *33*(2), 629-638.
4. Basera, C.H. *Journal of Education and e-Learning Research*, **2019**, *6*(2), 76-81.
5. Bhola, H.S. (Ed.). *Evaluating Functional Literacy*. London: Hulton, **1979**.
6. Black, P., Wiliam, D. *Assessment in Education: Principles, Policy and Practice*, **2018**, *25*(6), 551-575.
7. Broadbent, J., Panadero, E., Bouda, D. *Assessment and Evaluation in Higher Education*, **2018**, *43*(2), 307-322.
8. Buchholtz, N.F., Krosanke, N., Orschulik, A.B., Vorhölter, K. *ZDM – Mathematics Education*, **2018**, *50*(4), 715-728.
9. Carter, A., Bathmaker, A.-M. *Journal of Further and Higher Education*, **2017**, *41*(4), 460-474.
10. Cencelj, Z., Aberšek, M.K., Aberšek, B., Flogie, A. *Journal of Baltic Science*

Education, **2019**, *18*(1), 132-142.

11. Chipman, H.E., Rodríguez, F.J., Boyer, K.E. “I Impressed Myself with How Confident I Felt”: Reflections on a Computer Science Assessment for K-8 Teachers. *Proceedings of the 50th ACM Technical Symposium on Computer Science Education (pp. 1081-1087)*, Minneapolis, USA, **2019**.
12. Dolenc, K., Aberšek, B., Aberšek, M.K. *Journal of Baltic Science Education*, **2015**, *14*(2), 162-171.
13. Glogger, I., Holzäpfel, L., Kappich, J., Schwonke, R., Nückles, M., Renkl, A. *Education Research International*, **2013**, *2013*, article ID 785065.
14. Guo, W.Y., Yan, Z. *Assessment in Education: Principles, Policy and Practice*, **2019**, *26*(6), 675-699.
15. Haber, N., Mitchell, T.N. *Journal of Library and Information Services in Distance Learning*, **2017**, *11*(3-4), 300-313.
16. Haring, P., Warmelink, H., Valente, M., Roth Ch. *International Journal of Computer Games Technology*, **2018**, *October*, article ID 8784750.
17. Houston, D., Thompson, J.N. *Journal of University Teaching and Learning Practice*, **2017**, *14*(3), article number 2.
18. Kim, P.H. *Asia Life Sciences*, **2019**, *20*(2), 119-126.
19. Krasnoborova, A.A. *Criteria assessment as a means of improving the quality of education in the logic of the competency-based approach*. Perm: Perm State Pedagogical University, **2009**.
20. Leidinger, M., Perels, F. *Education Research International*, **2012**, *2012*, article ID 735790.
21. Lookadoo, K.L., Bostwick, E.N., Ralston, R., Elizondo, F.J., Wilson, S., Shaw, T.J., Jensen, M.L. *Journal of Science Education and Technology*, **2017**, *26*(6), 597-612.
22. Luckin, R., Clark, W., Avramides, K., Hunter, J., Oliver, M. *Interactive Learning Environments*, **2017**, *25*(1), 85-97.
23. Marinkovich, J., Sologuren, E., Shawky, M. *Circulo de Linguistica Aplicada a la Comunicacion*, **2018**, *74*, 195-220.
24. McGill, M.M, Decker, A.B. *CSeDResearch.org: Resources for primary and secondary computer science education research*. Aberdeen: University of Aberdeen, **2019**.
25. Meek, S.E.M., Blakemore, L., Marks, L. *Assessment and Evaluation in Higher Education*, **2017**, *42*(6), 1000-1013.
26. Methodological and educational-methodical foundations for the implementation of a system of criteria-based assessment of educational achievements of students. Toolkit. Astana: NAO named after I. Altynsarin, **2017**, https://nao.kz/blogs/download/117, accessed 25 December 2019.
27. Mirmotahari, O., Berg, Y., Fremstad, E., Damsa, C. *IEEE Global Engineering Education Conference, EDUCON*, **2019**, *April-2019*, 1152-1157.
28. Mohamadi Zenouzagh, Z. *Asia Pacific Education Review*, **2019**, *20*(3), 343-359.
29. Mohamadi, Z. *Studies in Educational Evaluation*, **2018**, *59*, 29-40.
30. Müller, W., Rebholz, S., Libbrecht, P. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, **2017**, *10108 LNCS*, 480-489.
31. Novikov, D.A. *Statistical methods in pedagogical research (typical cases)*. Moscow: MZ-Press, **2004**.
32. Ohashi, Y., Yamachi, H. Developing evaluation criteria for a service-learning course in computer science education. *Proceeding of the 5th International Conference on Information and Education Technology (pp. 53-57)*, Tokyo, Japan, **2017**.
33. Palmiero, C., Cecconi, L. *Journal of E-Learning and Knowledge Society*, **2019**, *15*(3), 89-99.
34. Pastore, S., Manuti, A., Scardigno, A.F. *European Journal of Teacher Education*, **2019**, *42*(3), 359-374.
35. Pieterse, V., du Toit, C.M. You asked for IT!: phrasing questions for computer science assessment, **2009**, https://dl.acm.org/doi/abs/10.1145/1562741.1562762, accessed 26 December 2019.
36. Psomos, P., Kordaki, M. *Proceedings of E-Learn*, **2011**, *37*(4), 17-21.

37. Rosário, P., Mourão, R., Trigo, L., Suárez, N., Fernández, E., Tuero-Herrero, E. *Psicothema*, **2011**, *23*(4), 681-687.

38. Sagimbaeva, A.E., Zaslavskaya, O.Yu., Avdarsol, S. *Criteria for assessing academic achievement in the Republic of Kazakhstan*. Moscow: Moscow State Pedagogical University, **2019**.

39. Sanchez, C.E., Atkinson, K.M., Koenka, A.C., Moshontz, H., Cooper, H. *Journal of Educational Psychology*, **2017**, *109*(8), 1049-1066.

40. Shalashova, M.M., Shevchenko, N.I., Mahotin, D.A. *Espacios*, **2018**, *39*(30), 1-13.

41. Shute, V.J., Hansen, E.G., Underwood, J.S., Razzouk, R.A *Education Research International*, **2011**, *2011*, article ID 915326.

42. Silseth, K., Gilje, Ø. *Assessment in Education: Principles, Policy and Practice*, **2019**, *26*(1), 26-42.

43. Sokolova, E.V. *Criteria internal assessment of educational achievements of students in grades 7-9 in teaching geometry*. Moscow: Moscow State Pedagogical University, **2017**.

44. Tienson-Tseng, H.L. *ACS Symposium Series*, **2019**, *1337*, 219-243.

45. Totikova, G., Aldabergenov, N., Salmirza, J., Nazarova, G., Gurbanova, A., Madiyarov, N., Yessaliyev, A. *Elementary Education Online*, **2019**, *18*(2), 461-471.

46. Vernon, A. *Innovative Higher Education*, **2011**, *26*(1), 59-77. DOI: 10.1023/A:1010938521365.

47. Veryaev, A., Nechunaeva, M., Tatarnikova, H. *Izvestiya of Altai State University*, **2012**, *2*(2), article number P13-17 2013.

48. Wang, Y., Li, H., Ding, Z. *Computational Intelligence and Neuroscience*, **2018**, *2018*, article ID 9745639.

49. Yang, S.-C., Luo, Y.-F., Chiang, C.-H. *Journal of Medical Internet Research*, **2017**, *19*(1), article number e15.

50. Yang, S.C., Luo, Y.F., Chiang, C.-H. *Journal of Medical Internet Research*, **2019**, *21*(11), article number e13140.

$$\bar{x} = \frac{1}{N}(x_1 + x_2 + x_3 + \ldots + x_{n-1} + x_n) = \frac{1}{N}\sum_{i=1}^{N} x_1 \quad \text{(Eq. 1)}$$

$$D_x = \frac{1}{N-1}\sum_{i=1}^{N}(x_1 - \bar{x})^2 \quad \text{(Eq. 2)}$$

$$T_{эмп} = \frac{\sqrt{M \cdot N}\,|\bar{x} - \bar{y}|}{\sqrt{M \cdot D_x + N \cdot D_y}} \quad \text{(Eq. 3)}$$

Table 1. *SAS assignments on the topic “Measuring information and computer memory”*

The objective of learning	7.1.3.1 Classification of computer networks
Criteria of assessment	Student: – determines the types of computer networks, information devices; – computer networks determine large-scale and local networks; – divides computer networks into wired and wireless networks; – is able to use types of networks in life; – types of networks, topology, information
Level of thinking skills	Use Skills of high level

Execution time	20 minutes

Table 2. *Test questions on classification of computer networks*

Criteria of assessment	Tasks No.	Descriptor Student	Grade
7.1.3.1 Classification of computer networks	1	Determines the types of computer networks, information devices	1
	2	Is able to write names of network topology	1
	3	Provides examples of using the topology of a network of schools, classes, buildings, institutions	1
	4	Identifies advantages and disadvantages of computer network	1
	5	Determines characteristics of definitions of local, big, global, zipping, text, table, regional, state, algorithm, package, image, format, wireless, cable, etc.	1
	6	In order to unite several cellphones, choses a type of network, defines in practice	1
		In total	6

Table 3. *The results of measurements of the level of knowledge in the control and experimental groups before and after the experiment*

No.	The control group (the number of correctly solved problems before the start of the experiment)	The experimental group (the number of correctly solved problems before the start of the experiment)	The control group (the number of correctly solved problems after the end of the experiment)	The experimental group (the number of correctly solved problems after the end of the experiment)
1	15	12	16	15
2	13	11	12	18
3	11	15	14	12
4	18	17	17	20
5	10	18	11	16
6	8	6	9	11
7	20	8	15	13
8	7	10	8	7
9	8	16	6	14
10	12	12	13	17
11	15	15	17	19
12	16	14	19	16
13	13	19	15	12
14	14	13	11	15
15	14	19	9	19
16	19		19	
17	7		8	
18	8		6	
19	11		9	
20	12		12	

Table 4. *The results of the experiment by level of knowledge*

The level of knowledge	Before the experiment		After the experiment	
	CG	EG	CG	EG
Low	6	3	7	1
Medium	10	8	8	7
High	4	4	5	7

Table 5. *The numbers of correctly solved problems in the control group before the experiment*

Parameters	Control group before the start of the experiment	Control group after the end of the experiment	Experimental group before the start of the experiment	Experimental group after the end of the experiment
Size of a sample	20	15	20	15
Minimum	7	6	6	7
Maximum	20	19	19	20
Interval (range)	13	13	13	13
Sum	251	205	246	224
Average	12.55	13.6667	12.3	14.9333
Median	12.5	14	12	15
Dispersion	15.31	15.24	16.75	12.49

Table 6. *Criterion Method of Cramer-Welch*

15	6.0025	15	0.004444
13	0.2025	18	9.404444
11	2.4025	12	8.604444
18	29.7025	20	25.67111
10	6.5025	16	1.137778
8	20.7025	11	15.47111
20	55.5025	13	3.737778
7	30.8025	7	62.93778
8	20.7025	14	0.871111
12	0.3025	17	4.271111
15	6.0025	19	16.53778
16	11.9025	16	1.137778
13	0.2025	12	8.604444
14	2.1025	15	0.004444
14	2.1025	19	16.53778
19	41.6025		
7	30.8025		
8	20.7025		
11	2.4025		
12	0.3025		
number	20		15
average	12.55		14.93333
dispersion	15.31316		12.49524
value of a criterion:	1.884972		

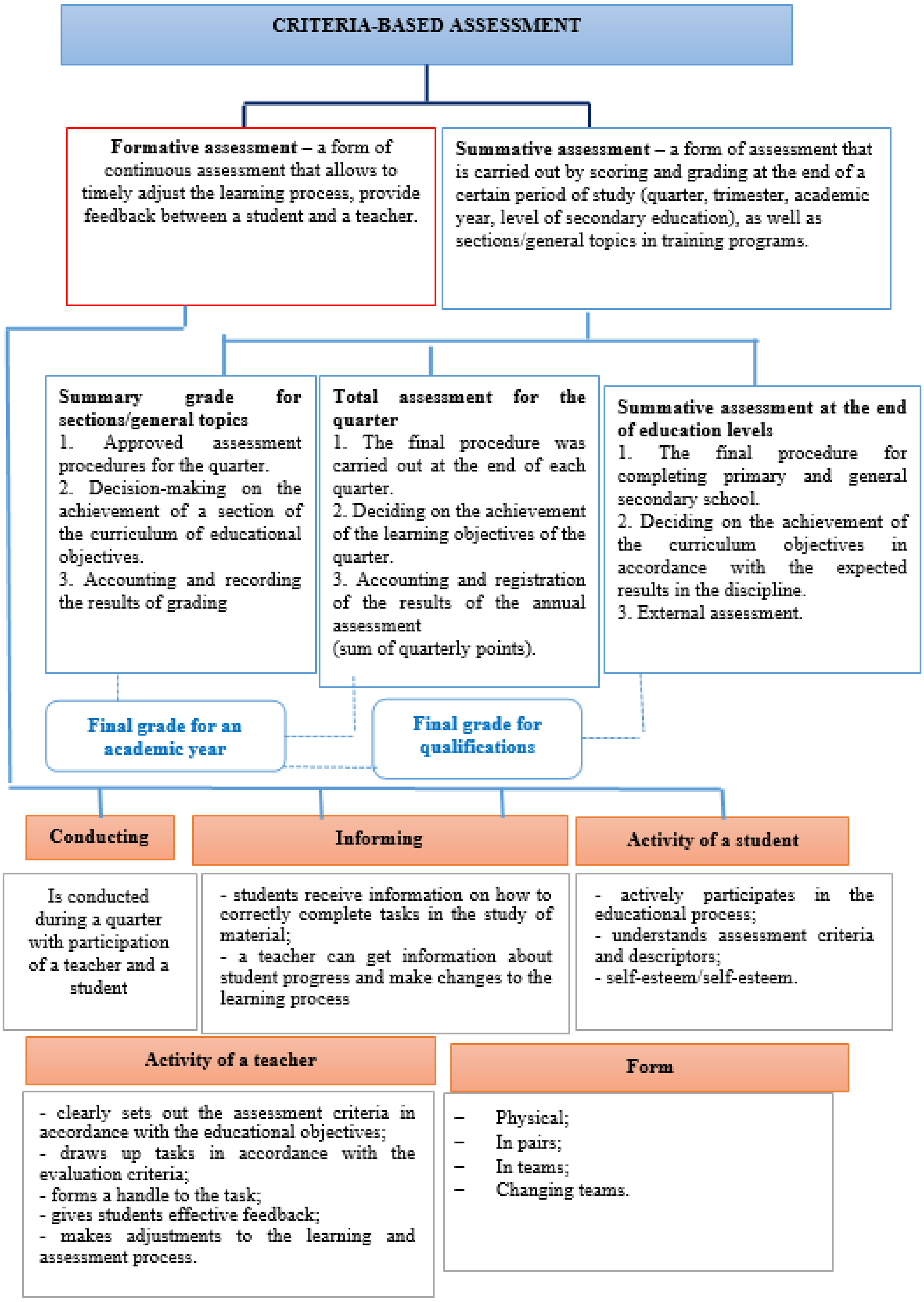

Figure 1. *The model of criteria-based assessment*
Reference: developed by the authors.

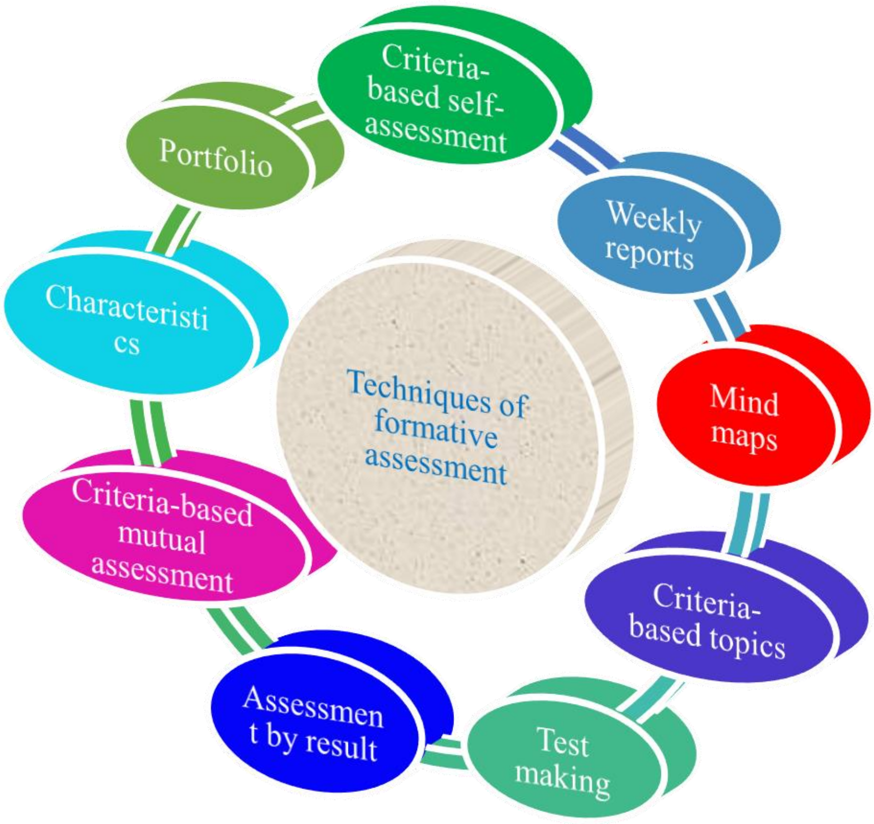

Figure 2. *Formative assessment methods*

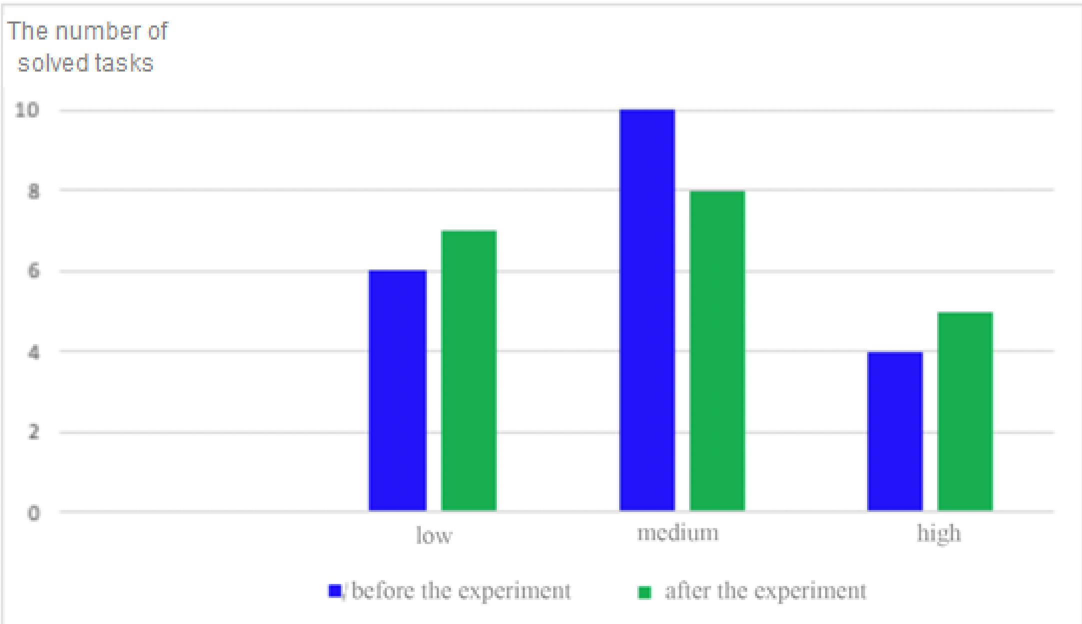

Figure 3. *Chart of knowledge levels*