Evaluation and Assessment in TEL courses

Pavel Boytchev[1,a)] and Svetla Boytcheva[2,b)]

[1]*Department of Information Technologies, Faculty of Mathematics and Informatics, Sofia University St. Kliment Ohridski, 5 James Bourchier blvd, Sofia 1164, Bulgaria*
[2]*Linguistic Modeling and Knowledge Processing Department, Institute of Information and Communication Technologies, Bulgarian Academy of Sciences, 25A, acad. G. Bonchev street, Sofia 1113, Bulgaria*

[a)]Corresponding author: boytchev@fmi.uni-sofia.bg
[b)]svetla.boytcheva@gmail.com

Abstract. This paper presents the evaluation in and the assessment of four university courses in which new technologies are widely used. Different factors of the evaluation are analyzed, along with their positive and negative aspects for both students and educators. It is investigated how the assessment of this evaluation is used to improve the courses themselves in two aspects - the included teaching material and in terms of students training approaches.

INTRODUCTION

Traditional Computer Science (CS) university-level courses are supported by several types of educational approaches and technologies, like Learning Management Systems, Distance learning, Web-based learning, e-Learning, Technology Enhanced Learning (TEL) and Virtual Learning Environments. Technology assisted learning of CS covers different levels of comprehension of the courses' material and plays a crucial role in development of skills. Active exploration, reflective learning, generalization, and the development of abstract thinking are the main objectives of Computer Graphics (CG) related courses in CS curricula, where the concepts are with higher abstraction. Additionally, a primary interest is the active exploration and generalization of these concepts. The main goal of CS courses is to develop abstract thinking rather than memorizing concepts and paradigms, thus the effectiveness of CS learning is achieved when the student is navigated and passes through all four stages of the learning cycle.

There is no globally accepted unique interpretation of the terms evaluation and assessment. In this paper, we use the interpretation[1] proposed by the Institute for Teaching, Learning and Academic Leadership at University of Albany, namely: *evaluation* is the measuring for the purpose of judging the value, while *assessment* or *formative assessment* is the measuring for the purpose of improvement. In this respect, the paper discusses evaluation of students, i.e. how their scores are formed; and assessment of the courses, i.e. how the courses are improved, based on the analysis of the evaluation of students.

This paper presents different evaluation metrics for four TEL CG courses. The evaluation includes computer based tests, oral exams, homework assignments, and course projects.

RELATED WORK

A standardized approach for effective assessments of TEL courses still does not exists, despite the various systematic studies and comprehensive frameworks developed in this area. The main reason for the lack of such standard is that assessments differ depending on the course subject, course objectives, level of education and national and institutional regulations for student evaluation. Booth et al [1] present four stages for designing a quality assessment for e-Learning: planning; developing; implementation and assessor support; and review. Cook and Ellaway present in [2] a comprehensive framework for TEL assessment in medical education. Cook describes three general approaches

[1]http://www.itlal.org/?q=node/93

of assessment: objectives-oriented; participant-oriented and process-oriented. A combination of these approaches is useful to understand complex interrelations between all three assessment approaches. Another common model for assessment is Context-Inputs-Processes-Products model described by Stufflebeam [3][2].

There is a lot of research in the field of e-Learning evaluation. Karran et al [4] based on the research in [1] and [5] describe advantages and disadvantages of commonly used e-Learning evaluation and feedback methods: online discussion, bulletin boards, collaborative assignments, self-assessment, online exams, online quizzes, computer-marked assignments, portfolios, role play, simulations, email, web publication, web design and development, and peer review.

EVALUATION IN TEL COURSES

This section presents the evaluation of students in four CG courses in the Faculty of Mathematics and Informatics at Sofia University *Fundamentals of Computer Graphics* (FCG), *Computer Graphics with WebGL* (CGW), *Educational Languages and Environments* (ELE) and *Geometry of Motion* (GM). These courses utilize TEL and rely on extensive use of modern technologies, which provide a suitable platform for demonstrating and experimenting with CG [6] [7]. These courses are picked for this review, as they are different in focus, complexity and approaches for student evaluation. Table 1 lists the components used for evaluation and their impact on the final grade.

TABLE 1. Count and impact of evaluation components in Computer Graphics related courses

Course	Homework	Test	Project	Exam	Bonus
FCG	2 homeworks low impact	3 tests high impact	high impact	high impact	low impact
CGW	11 homeworks high impact	-	high impact	-	low impact
ELE	-	4 tests high impact	high impact	-	low impact
GM	-	-	high impact	-	low impact

The evaluations are designed to consider several factors, and the most important one is the *complexity of the contents*. Complex courses require comprehensive evaluation – for example, the most difficult course is CGW and a student's final grade depends on 13 components. The second factor is the *number of students*. Each year FCG is taught to 120-140 undergraduates and their evaluation is designed to provide faster but still accurate scores. The last factor is the *orientation of the course*. Some courses are more aesthetically oriented, others are more technically oriented. This affects the criteria used in the evaluation components.

Although the student's evaluations in the courses are implemented differently, they are bound to the same frame of requirements. Namely, *over-excellent marks* (achieving scores higher than the maximal scores, more details are provided in the next sections), *multicomponent evaluation* (scores are calculated on performance in numerous activities), and *clarity* of the evaluation algorithm.

Fundamentals of Computer Graphics

FCG provides theoretical knowledge of fundamental concepts, algorithms and solutions in CG; as well as practical skills in programming 3D scenes – Fig. 1. The student's evaluation is calculated based on two homework assignments (5 pts each), two attendance tests (30 pts each), one remote test (5 pts), course project (60 pts), oral exam (40 pts) and bonuses (10 pts):

$$score = \left[\left[\sum_{i=1}^{3} test_i + \sum_{i=1}^{2} homework_i \right]_0^{60} + \max\left([project]_0^{40}, exam \right) + bonus \right]_0^{100} \tag{1}$$

where $[x]_a^b$ clamps x to the interval $[a,b]$. Over-excellent scores are supported in the evaluation of the tests, the project and the total score. For example, the total points for a project are 60, but they are clamped to [0,40], so students could reach the maximum of 40 even if their project is not perfect.

[2] [2] and [3] may use another interpretation of *evaluation* and *assessment* than the one used in this paper.

FIGURE 1. Students' projects in Fundamentals of Computer Graphics

Computer Graphics with WebGL

CGW is focused on learning and using WebGL – browser-based hardware-accelerated 3D graphics. The educational content is considerably more complex. To maintain consistent learning curve, the students' evaluation is based on 11 weekly homework assignments, a course project and bonuses:

$$score = \frac{1}{10}\left(\sum_{i=1}^{11} homework_i + project + bonus\right) \tag{2}$$

The evaluation of the project is based on over-excellent scoring of the following criteria: complexity of motion, complexity of objects, graphical effects and interactivity. The top score could be reached by covering just two of these four criteria. This provides flexibility to students, as some of them would prefer to make a complex 3D scene with special effects and no motion, while others would prefer to program complex and interactive animation – Fig. 2.

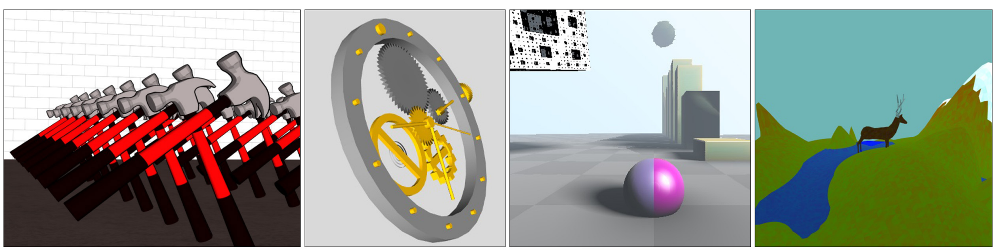

FIGURE 2. Students' projects in Computer Graphics with WebGL

Educational Languages and Environments

The ELE course uses a specially developed virtual environment called *Suica*. Students learn how to design and build interactive graphical educational content – Fig. 3. As a programming complexity, this course is lighter than FCG and CGW, as it is targeted towards future teachers, not towards CS experts. The evaluation is based on four tests, a course project and bonuses and the overall score is calculated rather straightforward as the sum of the individual scores clamped to [0,100]:

$$score = \left[\sum_{i=1}^{4} test_i + project + bonus\right]_0^{100} \tag{3}$$

The structures of the tests and the project are more complex. Each test consists of 20 practical micro-problems, and the course project is to creste a TEL-enabled lesson, evaluated in respect to 10 criteria about the artistic, pedagogical and technological merits.

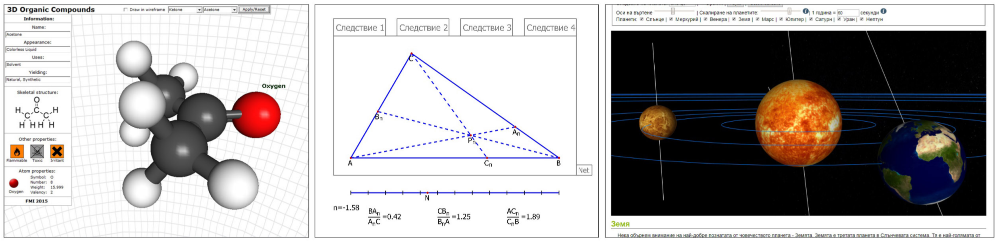

FIGURE 3. Students' projects in Educational Languages and Environments

Geometry of Motion

GM is a multidisciplinary course blending Physics, Mathematics and Computer Science [8]. The GM lab exercises develop skills to programmatically design and animate virtual mechanisms [9] [10] – see Fig. 4 – in a virtual environment called *Mecho* (MECHanical Objects) [11]. Apart from the oral exam, the students have to build their own virtual mechanisms [7]. The score is based on an oral exam and a project, evaluated against 20 criteria in 4 groups: administrative, visual, hardware and software experience. The evaluation process adopts the waterfall model – each group of criteria is evaluated, only if at least 3 criteria of the previous group are fulfilled:

$$score = K_0 + project = K_0 + \frac{5}{3}\left(\left[\sum_{i=1}^{4} K_i + bonus\right]_0^{20} - 5\right) \tag{4}$$

where K_0 is the score of the oral exam and $K_{i=1..4}$ are the scores of the i-th group of criteria.

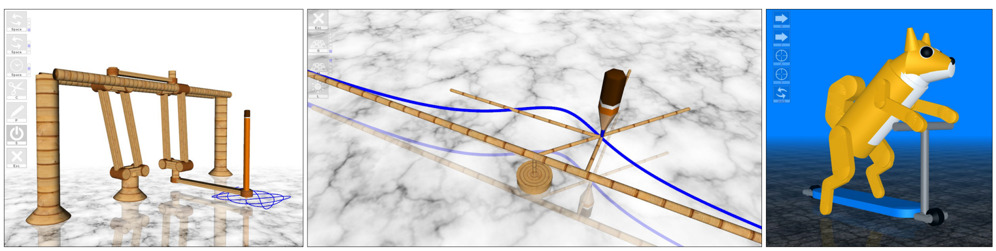

FIGURE 4. Students' projects in Geometry of Motion

DISCUSSION OF THE EVALUATION

The design of the scoring in the presented TEL courses and their actual implementations have both advantages and disadvantages. This section presents the main features of the evaluation and their impact on the student and the lecturer. The result of this analysis is used to form the factors for Tables 2 and 3.

All presented courses include the submission of a software project. Observations show that most of the submissions were made during the last day before the exam, so the deadline was shifted to a week before the exam and the students are given an option to correct their projects after their submission, but before the exam.

As for homework assignments and tests, most students do prefer to have their scores as early as it is possible. The evaluation procedure is designed in a way to provide feedback within a few hours. This is achieved by restructuring the tests' and homework assignments' content. Before the beginning of each academic year all tests and homeworks are reviewed. Test questions which appear to be confusing or unclear to the majority of the students, or require substantial efforts to be evaluated, are being modified or replaced.

The score for each course is calculated on the students' performance in several activities, which are all optional. Students may select to ignore some activities, if they could secure sufficient points from other activities. Many students, especially those who do not aim at high results, welcome this. The downside is that the final score of some students is lower than what they could actually achieve.

The evaluation in all courses may generate over-excellent marks. This is achieved by a grading scale with maximum below the actual maximum. The introduction of over-excellent scores is beneficial to students, especially in evaluations with diverse criteria. Students could focus on some of the criteria, and provide elegant solutions only for them. This allows students with mathematical interests and other students with artistic interests to show both excellent performance and results.

All four courses are managed online – tests are performed online, homework assignments are submitted online, the gradebook is available online, forums and chatrooms allow online communication. This fosters immediate interaction and feedback, easier searching and filtering and more ubiquitous learning experience. The most significant impact of online teaching management is the execution of online tests – monitoring restrictions (passwords, IP addresses, time slots), providing rich content, on-the-fly generation of tests and automatic evaluation of closed questions.

The use of LMS for managing educational activities is advantageous and disadvantageous at the same time. It is not possible to present any aspect of TEL teaching or learning as only positive, or only negative. At the end of every semester the students take part in anonymous polls and provide feedback about the course content, the evaluation criteria in the course and the teaching – see Fig. 5. Each poll contains questions in five categories, related to the course as a whole, the professor, the teaching assistant (TA), the course content. The poll also contains options for general feedback and recommendations in free text format.

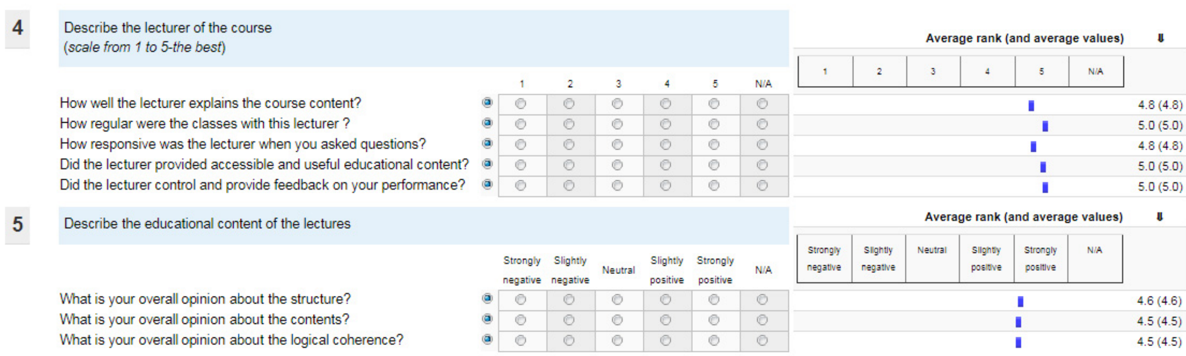

FIGURE 5. Fragment of the online poll for FCG in the winter semester of 2017

Professors have access to the results of the poll in digital form. The most informative part of the poll is the pros and cons section where students discuss freely what they like or don't like in the course, including their opinions about the evaluation procedures. Based on feedback results of 400+ students, a list of nine factors related to the evaluation approaches used in the courses and their impact are presented in Table 2 and Table 3. The factors are neither positive only, nor negative only. Moreover, it is a matter of appropriate balance in order to guide each factor to express its advantageous features. The results in Table 3 are based on authors' personal experience of over 10 years of teaching TEL courses. Opinions of other faculty members are not included, as they are not taking part in the evaluations discussed in this paper.

TABLE 2. Evaluation approach impact - student's point of view

Factor	Advantages	Disadvantages
A new projects' deadline	A chance to correct submitted projects	Collision with other exams
Fast evaluation of tests	Quick feedback with results	Getting preliminary information
Optional activities	Sparing efforts	Underestimated scores
Exemption from project	Stimulus for better performance at tests	Disappointment if score is insufficient
Over-excellent score	Easier to achieve good marks	Spending more efforts
On-line content	No problems if classes are skipped	Some content could not be online
On-line gradebook	Monitoring current progress	Layout cannot be customized
On-line submission	Unified interface and notifications	Occasional technical issues
On-line tests	Questions with animations	Occasional technical issues

TABLE 3. Evaluation approach impact - lecturer's point of view

Factor	Advantages	Disadvantages
A new projects' deadline	Sufficient time for project evaluation	Exam activities spread over two days
Fast evaluation of tests	Better distribution of load	Getting preliminary information
Optional activities	Skipped activities are not evaluated	Annual readjustment of scores
Exemption from project	Stimulus for better performance at tests	Some disappointed students
Over-excellent score	Easier identification of potential TAs	Underestimation of over-excellent score
On-line content	Could be reused and referenced	Need for regular update
On-line gradebook	Automatic score calculation	Hard to customize the layout
On-line assignment	Automatic restrictions and notifications	Confusion with draft submissions
On-line tests	Richer questions, restrictions, logs	Easier cheating, need for training TAs

ASSESSMENT OF EVALUATION AND CONTENTS

According to the definitions of *evaluation* and *assessment* stated in the second paragraph of the introduction, *assessment of evaluation* is considered the process of measuring the evaluation of students in order to improve that evaluation. As described further in this section, the assessment of the evaluation contributes to the assessment of the course content. The evaluations presented so far are results of numerous fine-tunings over the past decade. The very first evaluations were plain flat scales with minimal number of evaluation components, which would be sufficient for non-TEL courses. However, such simplistic scoring of TEL courses did not provide enough accuracy and finer precision of students' results.

Table 4 and Table 5 show the students' final grades in two of the courses (FCG and ELE) as distribution of A's, B's, C's, D's and F's; the bottom line is the total number of students. Figure 6 visualizes these distributions over the years with aggregated data for B's, C's and D's. FCG experienced two major redesigns – initially it used *Elica* [12] and [13], then *FMI3D* (a virtual environment written in C++ and OpenGL) and currently it uses *Meiro* [14] (written in JavaScript and Three.js). According to Fig. 6, the evaluation results of students improved significantly after the FCG content was shifted from C++ to JavaScript (JS) in 2015 – the percentage of A's increased over 50%, while the number of F's remained relatively unchanged.

TABLE 4. Students' evaluation results in FCG from 2010 to 2017

Score	2012	2013	2014	2015	2016	2017
A	28%	37%	40%	61%	53%	59%
B	20%	12%	4%	8%	6%	2%
C	27%	22%	21%	6%	9%	15%
D	12%	28%	26%	19%	29%	20%
F	13%	1%	9%	6%	2%	4%
Students	138	138	141	128	129	124

TABLE 5. Students' evaluation results in ELE from 2007 to 2017

Score	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
A	42%	69%	35%	56%	38%	23%	13%	13%	11%	22%	14%
B	16%	7%	19%	18%	6%	10%	4%	3%	18%	9%	0%
C	16%	11%	14%	18%	16%	15%	12%	14%	20%	14%	17%
D	21%	11%	5%	6%	3%	11%	12%	23%	34%	28%	38%
F	5%	2%	26%	3%	38%	42%	60%	46%	18%	27%	31%
Students	19	45	57	34	32	62	52	69	74	81	42

A more interesting observation could be made for the dynamics of the evaluation results for ELE. Considering the curve for F's, in 2007 and 2008 it was close to zero. The introduction of projects into the students' evaluation raises the F's to 26% in 2009. The amount of F's drops below 3% in 2010, as students get accustomed to the projects.

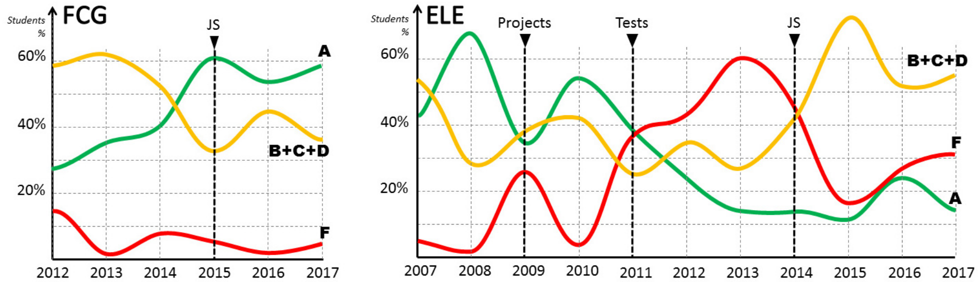

FIGURE 6. Students' evaluation results in FCG and ELE

Then in 2011 the tests are introduced and the percentage of F's raises up to 38%. The next two years there is a steady increase to 42% and then to 60%. This tendency prompted a redesign of the ELE contents and shifting it from *Elica* to *Suica+JS*. As a result, the percentage of F's drops down and stabilizes to the acceptable 30% for the next 3 years. This whole process demonstrates how the assessment of evaluation improves the both the evaluation process and the course contents.

Actually, not just FCG and ELE changed their evaluation procedure and course contents. As a result of the regular assessment of all courses, they are updated annually. Three of the four courses, however, experienced radical changes as indicated as break-lines in the historical development in Fig. 7. ELE experienced one such change in 2014, while FCG was significantly modified in 2013 and 2015.

CGW is a new course and only minor updates have been introduced to it for the last 4 years; however, it is a successor of Computer Graphics course (not discussed in this paper) [14]. CG and CGW are so radically different, that it was decided that CG will cease to exist in 2014, and CGW will replace it as a completely new course.

GM has the most active history. It started being based on *Elica*. When the first version of *Mecho* appeared (it was in *Elica*), the course shifted to *Mecho*. In 2011 Mecho was rewritten in C++ with OpenGL and the course changed again. Finally, in 2014 the new version of Mecho in JavaScript with WebGL led to the latest redesign of the course.

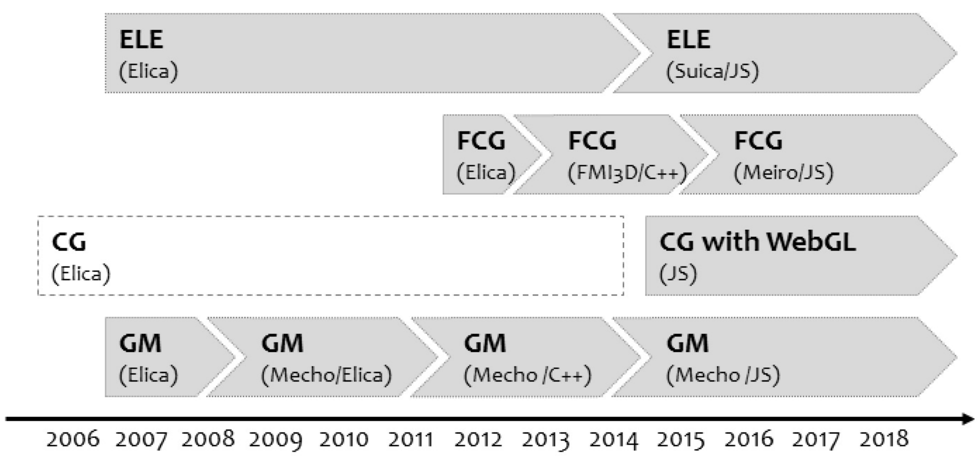

FIGURE 7. Historical development of ELE, FCG, CGW and GM courses

CONCLUSION AND FURTHER WORK

The proposed different evaluation criteria for four CG courses provide both flexibility and accuracy in the evaluation and the assessment in TEL. The evaluation techniques allow differentiating between levels of concept abstraction. Moreover, they provide a means for measuring not only student skills in CS and CG, but also help the evaluation of some interdisciplinary transitions and students' aesthetic perception.

The evaluation in the courses is evolving together with the change of their educational content. Although these courses did not change a lot for the past 3 years, there are plans for further improvement of the evaluation and the assessment. Namely, there are ongoing efforts to expand the virtual environment *Meiro* used in FCG to serve as an evaluation and assessment platform. *Meiro* is a game-like online virtual environment where students navigate in a multilayer 3D maze [15]. There are interactive models placed throughout the maze that allow the students to experiment with concepts from the domain of Computer Graphics.

Currently there are several hundreds 3D models in *Meiro* and the plans are to add models for student evaluation. From students' perspective, *Meiro* replaces formal tests will interactive 3D game. This is expected to increase the motivation, to eliminate the exam fright, to make the cheating meaningless and to merge learning and examining activities into one single activity. The practical benefits of such gamified evaluation will be the ubiquitous learning – students could play the game anytime, anywhere and anyhow (i.e. using different hardware devices and software platforms). The models for student evaluation will be based on prototypes, so for each student *Meiro* will generate a different model. Models will have mechanism to calculate a numerical score for the performance of each student. This score will contribute to the overall evaluation. As for the assessment, an aggregated analysis of the results from these models will provide insights about problematic sections from the course content, which will be addressed either by modifying the course, or by modifying the teaching.

Another possible further development of the students' evaluation and the formative assessment of these courses may be inspired by a new course, which is currently being developed. It is focused on Virtual and Augmented Reality (VR/AR). After the pilot run of this course, which is scheduled for the spring semester in 2019, it will be decided how the VR/AR technology could be embedded in the other TEL courses and in *Suica* and *Meiro*. This may again require a radical change in these courses.

REFERENCES

[1] R. Booth, B. Clayton, R. Hartcher, S. Hungar, P. Hyde, and P. Wilson, *The Development of Quality Online Assessment in Vocational Education and Training. Volume 1 [and] Volume 2.* (ERIC, 2003).

[2] D. A. Cook and R. H. Ellaway, Medical teacher **37**, 961–970 (2015).

[3] D. L. Stufflebeam, in *International handbook of educational evaluation* (Springer, 2003), pp. 31–62.

[4] T. Karran, in *eReflections – Ten years of Educational Technology Studies at the University of Oulu* (Oulu University Press, 2004), pp. 127–152.

[5] S. Kerka and M. E. Wonacott, Assessing learners online: Practitioner file, 2000.

[6] P. Boytchev, K. Kanev, and R. Nikolov, "Technology enhanced learning with subject field multiplicity support," in *Proceedings of the 2012 Joint International Conference on Human-Centered Computer Environments* (ACM, 2012), pp. 39–44.

[7] P. Boytchev and P. Armyanov, "Re–experiencing engineering inventions within a modern virtual environment," in *Proceedings of 2nd International Conference Software, Services & Semantic Technologies* (2010), pp. 55–62.

[8] P. Boytchev, "Equilibristic pandisciplinary approach to technology enhanced learning. mathematics and mathematical education," in *Proceedings of the 40th Spring Conference of the Union of Mathematicians in Bulgaria* (2011), pp. 340–346.

[9] P. Boytchev, in *Research on e-Learning and ICT in Education* (Springer, 2012), pp. 267–281.

[10] P. Boytchev, E. Sendova, and E. Kovatcheva, International Journal on Information Technologies and Security (IJITS), Year III 27–40 (2011).

[11] P. Boytchev, "Mecho - educational software for virtual mathematical devices," in *Proceedings of 3rd EDUvision International Conference, EDUvision* (2013), pp. 692–707.

[12] P. Boytchev, "Natural Object-oriented programming: The evolution metaphor behind Elica," in *Proceedings of the 2 nd ISSEP International Conference, eds.: Dagiene V., Mittermeir R* (2006), pp. 351–360.

[13] P. Boytchev, "The niche for educational software in mathematical learning," in *International Conference, Pioneers of Bulgarian Mathematics, Technology Enhanced Learning section, Sofia, Bulgaria* (Citeseer, 2006).

[14] P. Boytchev, Informatics in Education-International Journal **6**, 269–282 (2007).

[15] M. Lekova and P. Boytchev, "Virtual learning environment for computer graphics university course," in *Proceeding of 12th International Technology, Education and Development Conference,IATED Academy* (2018), pp. 3301–3309.