Information Channel Based Measure of Effectiveness of Computer-Assisted Assessment in Flipped Classroom

Jerzy Rutkowski(✉) and Katarzyna Moscinska

Akademicka 16, 44-100 Gliwice, Poland
{jrutkowski,kmoscinska}@polsl.pl

Abstract. Properly designed Assessment Program, both Formative tests and Summative test (Exam), is the basic part of arbitrary engineering course. Today, computerization of Assessment is a norm and in the 2nd generation of Smart Classroom, computerization means first of all active use of mobile technology. Then, measurement of effectiveness of Computer-Assisted Assessment, achievements of learning outcomes, correlation between number of Formative tests solved by the students and the Exam Pass/Fail ratio, is absolutely essential. The proposed measurement method is based on Discrete Memoryless Channel principles and utilizes mutual information as the measure of this correlation. The Case Study is presented and it confirms usefulness of the proposed measure. Some guidelines, good practices in Computer Assisted Assessment design are given.

Keywords: Technology enhanced learning · Computer Assisted Assessment

1 Introduction

In the era of dynamic development of ICT, their use in Higher Education is ubiquitous. The ICT, properly used, may significantly contribute to the quality of education, enable development of new course delivery methods, new methods of assessment. Learning Activities are the core of every engineering course and these Activities can be divided into two components:

- Learning Content (knowledge delivery);
- Assessment Program (knowledge assessment).

These components have to be tightly correlated, regardless the course delivery method, in the Flipped Classroom [1] as well as in the Traditional Classroom. In the Flipped Classroom, the Learning Content is based on e-materials. Then, Assessment Program should be correlated with these e-materials and take into account the Flipped Classroom characteristics, Learning Content modular structure in particular. Today, computerization of Assessment in HE is a norm. Tests (quizzes) are completed by the student at a computer, firstly at home (Formative Assessment), then in a computer lab (Summative Assessment, Exam), without the teacher's intervention. Extensive research in Computer-Assisted Assessment (CAA), also called E-Assessment and Computer-Aided Assessment, has been done so far, its effectiveness has been reported [2–4].

V.L. Uskov et al. (eds.), *Smart Education and e-Learning 2017*, Smart Innovation, Systems and Technologies 75, DOI 10.1007/978-3-319-59451-4_2

However, further studies are necessary to make the Assessment Program fully computer-automated and reliable, such that traditional Assessment can be replaced by CAA. In fact, traditional Assessment has to be practically eliminated from the Flipped Classroom. Organization of Moodle-based CAA in the Flipped Classroom has been discussed in [5]. To make this CAA effective, students have to accept new learning model, teachers have to match their teaching to this model:

- Students have to accept Self-Regulated Learning (SRL) mode of learning, take Formative tests systematically and solve questions with understanding.
- Teachers have to prepare high quality e-materials supporting the SRL. Guidelines how to prepare good e-materials, using different techniques, have been given by many authors, the review can be found in [1]. Then, to ensure high correlation between the number of Formative questions answered and the Exam grade, the teacher has to properly construct and align the Formative tests and the Exam test. Moreover, the teacher has to take into account that, in the 2nd generation of Smart Flipped Classroom, implementations are mainly based on active use of mobile technology and automatic communications in the Smart Classroom environment [6]. Obviously, when designing the CAA, achievements of learning outcomes have to be taken into account as well.

A model of SRL, also called Self-Directed Learning (SDL), correlation between student goals, tactics, strategies and achievements of learning outcomes have been discussed by many authors [7, 8] – this model has been repeated in Fig. 1.

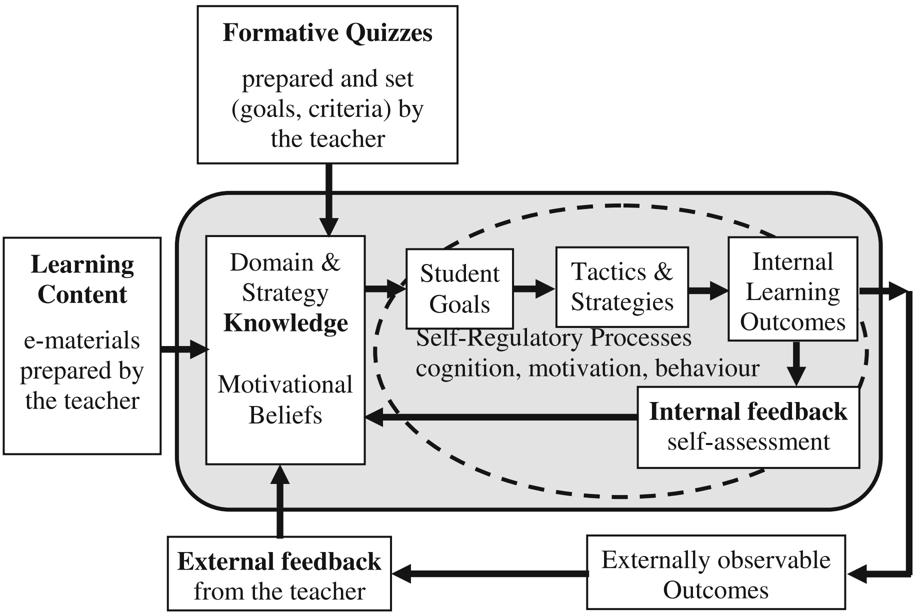

Fig. 1. Model of self-regulated learning

While there would normally be an overlap between the student goals and those of the teacher, the degree of overlap may not be high. The following main barriers in making Formative Quizzes more effective can be enlisted:

1. Students' SDL Readiness (SDLR) [9] is low, freshmen SDLR in particular.
2. If students perceive Formative assessment as primarily examining content knowledge, they will tend to do little more than rote learning, especially when they wish only to pass the Exam [10].
3. Students tend to ignore activities that do not directly contribute to grades and degree class; even though they could see the benefit of developing competencies, they do not take advantage of it [10].

These barriers can be broken by providing students with clear evidence that correlation between Formative quizzes taken and the Exam result is very high. Solving of Formative quizzes is supported by e-materials provided by the teacher and obviously they are only as good as the teacher who prepares them. Then, measurement of correlation between the number of Formative questions solved and the Exam score is essential for both the teacher and the students. A new method for finding this correlation has been proposed in [8]. This method is based in the field of Information Theory [11], developed by C.E. Shannon in the late 40's of the last century. The updated version of this method is presented in Sect. 2. A case study, use of Information Theory to evaluate effectiveness of Formative quizzes for Electric Circuit Analysis course in the academic year 2015/2016, is presented in Sect. 3, some guidelines and conclusions are given in Sect. 4.

2 Evaluation of Correlation Between Formative Quizzes Taken and Exam Results

The relationship between activity in Formative Quizzes and Exam results can be described by means of Discrete Memoryless Information Channel (DMC). Figure 2 presents Source-Channel-Receiver information system [9], where:

$$\mathbf{X} = \{X_1, \ldots, X_M\} \tag{1}$$

is the discrete input source of information, in short the Source, set of samples (ensembles) characterized by the probability assignment

$$\mathbf{P_X} = \{p\{X_1\}, \ldots, p(X_M)\} \tag{1a}$$

$$\mathbf{Y} = \{Y_1, \ldots, Y_K\} \tag{2}$$

is the channel output source, in short the Receiver, characterized by the probability assignment

$$\mathbf{P_Y} = \{p\{Y_1\}, \ldots, p(Y_K)\} \tag{2a}$$

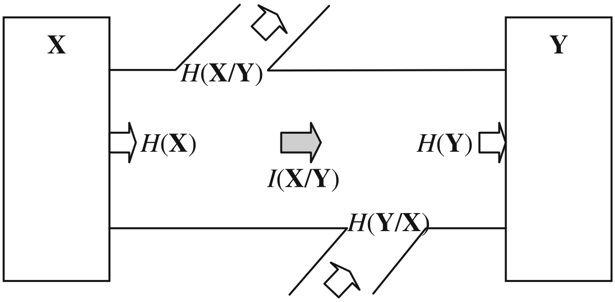

Fig. 2. Information system: source-channel-receiver

and Channel itself is characterized by *MK* transition probabilities that relate ensembles of input and output source:

$$p\left(Y_j/X_i\right); i = 1, \ldots, M, j = 1, \ldots, K \tag{3}$$

Then, for the given probabilistic model of the Source and the Channel, information loss $H(\mathbf{X/Y})$, misinformation $H(\mathbf{Y/X})$ and mutual information $I(\mathbf{X/Y})$ can be defined. Mutual information between events (sources) **X** and **Y** is the information provided about the event **X** by the occurrence of the event **Y**, or vice versa.

$$I\,(\mathbf{X}/\mathbf{Y}) = H(\mathbf{X}) - H(\mathbf{X}/\mathbf{Y}) = H(\mathbf{Y}) - H(\mathbf{Y}/\mathbf{X}) \tag{4}$$

$$H(\mathbf{X}) = -\sum_{i=1}^{M} p(X_i) \log_2 p(X_i) \tag{5a}$$

$$H(\mathbf{Y}) = -\sum_{j=1}^{K} p(Y_j) \log_2 p(Y_j) \tag{5b}$$

$$H(\mathbf{Y}/\mathbf{X}) = -\sum_{i=1}^{M} \sum_{j=1}^{K} p(X_i, Y_j) \log_2 p(Y_j/X_i) \tag{6}$$

To give a measure, how far the considered channel is from the idealized (target) one, the normalized mutual information is introduced:

$$I_{\mathrm{n}} = I(\mathbf{X}/\mathbf{Y})/I(\mathbf{X}/\mathbf{Y})_{\mathrm{ref}} \tag{7}$$

where $I(\mathbf{X/Y})_{\mathrm{ref}}$ is the mutual information of the reference channel, channel that is considered as the idealized target one.

Exemplary channels: an arbitrary binary channel and the reference 3-input/3-output channel, are presented in Fig. 3.

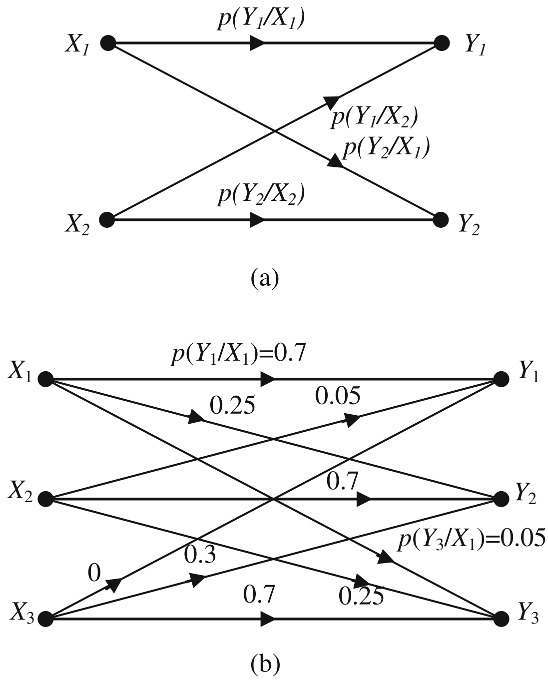

Fig. 3. (a) Exemplary channels: binary (b) Exemplary channels: 3-input/3-output (reference of Case Study)

Exam can be considered as measurement of students' knowledge and then, it can be described by Discrete Memoryless Information Channel (DMC). Students' knowledge, the measured quantity, can be expressed by a number of Formative quizzes taken and, after discretization, it consists the set of samples, the input Source. Set of Exam results consists the output source (Receiver). Exam can be considered as measurement of students' knowledge. It can be assumed that this knowledge is designated by Formative quizzes taken, number of tasks solved by the student. For the binary channel ($M = K = 2$), both sources can be discretized as follows:

- $X_1 = X_D$ Diligent students, D students that solved at least T_D % of tasks,
- $X_2 = X_N$ Negligent students, N students that solved less than T_D % of tasks,
- $Y_1 = Y_P$ P students that Passed Exam,
- $Y_2 = Y_F$ F students that Failed Exam.

It is assumed that numbers D, N, P, F are known and they designate probability assignments $\mathbf{P_X}$ and $\mathbf{P_Y}$, e.g. $p(X_D) = D/M$ is probability of Diligence, $p(Y_F) = F/M$ is probability of Fail, $M = D+N = P+F$. Also, conditional probabilities that relate students' Diligence/Negligence and Exam results: $p(Y_j/X_i)$, i = D,N; j = P,F, are known, e.g. $p(Y_P/X_N) = P_N/N$ is the probability of Passing the Exam by the Negligent student, where P_N is the number of Negligent students that Passed. Then, relationship between

Formative quizzes and the Exam can be modeled by means of binary information channel, as depicted in Fig. 3a, and mutual information can be calculated. This information may be interpreted as the information provided about the measured data (students' knowledge) by the occurrence of measurements (Exam), in other words as the measure of effectiveness of Formative quizzes.

To describe more precisely relationship between students activity in solving quizzes and their performance during the Exam, more complex channel can be considered. The ensemble of Diligent students can be split into Diligent High and Diligent Low, while the ensemble of Negligent students remains unchanged:

- $X_1 = X_{DH}$ D_H Diligent High students that solved more than T_{DH} % of tasks,
- $X_2 = X_{DL}$ D_L Diligent Low students that solved between T_{DL} % and T_{DH} %,
- $X_3 = X_N$ N Negligent students that solved less than T_{DL} % of tasks.

Then, Pass can be split into Pass High and Pass Low while the ensemble of students that Failed remains unchanged:

- $Y_1 = Y_{PH}$ P_H students that Passed with Excellent, Very Good or Good grade,
- $Y_2 = Y_{PL}$ P_L students that Passed with Satisfactory or Sufficient grade,
- $Y_3 = Y_F$ F students that Failed Exam.

The conditional probabilities have to be split accordingly, taking into account how many Diligent High Passed High, etc., e.g. $p(Y_{PH}/X_{DL}) = P_{H,DL}/D_L$ is the probability of Passing High by Diligent Low, where $P_{H,DL}$ is the number of Diligent Low that Passed High. Then, the relationship between Formative quizzes and Exam can be modeled by means of 3-input/3-output information channel. The conditional probabilities can be expressed in the form of table, as presented in Table 1.

Table 1. Conditional probabilities of 3-input/3-output Channel

$p(Y_j/X_i)$	$Y_1 = Y_{PH}$	$Y_2 = Y_{PL}$	$Y_3 = Y_F$
$X_1 = X_{DH}$	$P_{H,DH}/D_H$	$P_{L,DH}/D_H$	F_{DH}/D_H
$X_2 = X_{DL}$	$P_{H,DL}/D_L$	$P_{L,DL}/D_L$	F_{DL}/D_L
$X_3 = X_N$	$P_{H,N}/N$	$P_{L,N}/N$	F_N/N

3 Case Study: Electric Circuit Analysis Course

For the first time, the described methodology of evaluation of relationship between number of Formative quizzes taken and Exam results has been verified in the academic year 2014/2015. Fifty students (all enrolled) of Macro (Electronics + Automatics + Informatics), consisted the test group, results of only the first Exam have been taken into account. Formative quizzes and the Exam have been distributed through Moodle LMS. The obligatory Formative quizzes contained $9 \times 15 = 135$ *Calculated* questions [12] and the following thresholds have been assumed: obligatory minimum $T_N = 30\% = 40$, $T_{DL} = 50\% = 68$, $T_{DH} = 67\% = 90$ questions. The Exam quiz consisted of ten questions: eight *Calculated* questions, marked 0 or 1 and drawn from Formative quizzes, two *Multiple-Choice*, marked 1, 0 (no answer) or −0.5 (wrong answer). The Pass threshold at 3.5 points has been experienced as the most adequate [13]. This case study has been presented in [8], high correlation between students activity in solving Formative quizzes and the Exam results has been confirmed and the following main conclusions have been drawn:

1. Only students that have solved majority of Formative Quizzes (Diligent High students), Pass the Exam with High mark.
2. Minority of Negligent students that have solved only the obligatory number of T_N questions Passed, probability of Passing with a High mark was practically zero.
3. Very small percentage of Diligent High students Failed the Exam.

In the academic year 2015/2016 the study has been repeated, the test group consisted of 276 students (all enrolled), representing three fields of study: Macro (53), Informatics (158) and Teleinformatics (65). The set of Formative questions has been enlarged to $9 \times 21 + 100 = 289$ *Calculated* questions (number of questions per obligatory quiz has been increased from 15 to 21, nonobligatory quiz with 100 questions has been added) and the following new thresholds have been assumed: $T_N = 20\% = 55$, $T_{DL} = 33\% = 96$, $T_{DH} = 50\% = 145$ questions. The Exam organization remained unchanged, i.e. same Pass threshold of 3.5 points has been applied.

The results of the first two Exams have been taken into account, i.e. if the student failed the first Exam, then the resit result has been taken into account. The 3-input/3-output channel provides more information than the binary and only such channel has been considered. This channel, conditional probabilities $\mathbf{P_{Y/X}}$, together with the input probability assignment $\mathbf{P_X}$ and the output probability assignment $\mathbf{P_Y}$ (calculated from $\mathbf{P_X}$ and $\mathbf{P_{Y/X}}$), for each field of study and the aggregate, are presented in Fig. 4. Conditional probabilities for the aggregate channel are repeated in Table 2.

From formulas (4), (5a), (5b) and (6) the mutual information $I(\mathbf{X/Y})$ can be easily calculated and for the aggregate channel $I(\mathbf{X/Y}) = 0.40$bit. In an ideal channel (noiseless channel, all crossover probabilities are zero) $I(\mathbf{X/Y})_{max} = H(\mathbf{X}) = 1.30$bit but such idealization is too rigorous to give the reference channel. To get more realistic reference (target) channel, we may accept that 5% of Diligent High students Fail (explanation will be given in Sect. 4). We may accept even higher percentages for other crossover transitions except one, we may not accept that Negligent students Pass High, i.e. for the reference channel we have to assume $p(Y_{PH}/X_N)_{ref} = 0$. If we assume the reference (target) channel as proposed in Fig. 3b, then the reference mutual information $I(\mathbf{X/Y})_{ref} = 0.41$bit and consequently $I_n \approx 1$, $I_n\% \approx 100\%$.

This maximum normalized mutual information proves again the extremely high correlation between the Formative quizzes taken and the Exam results.

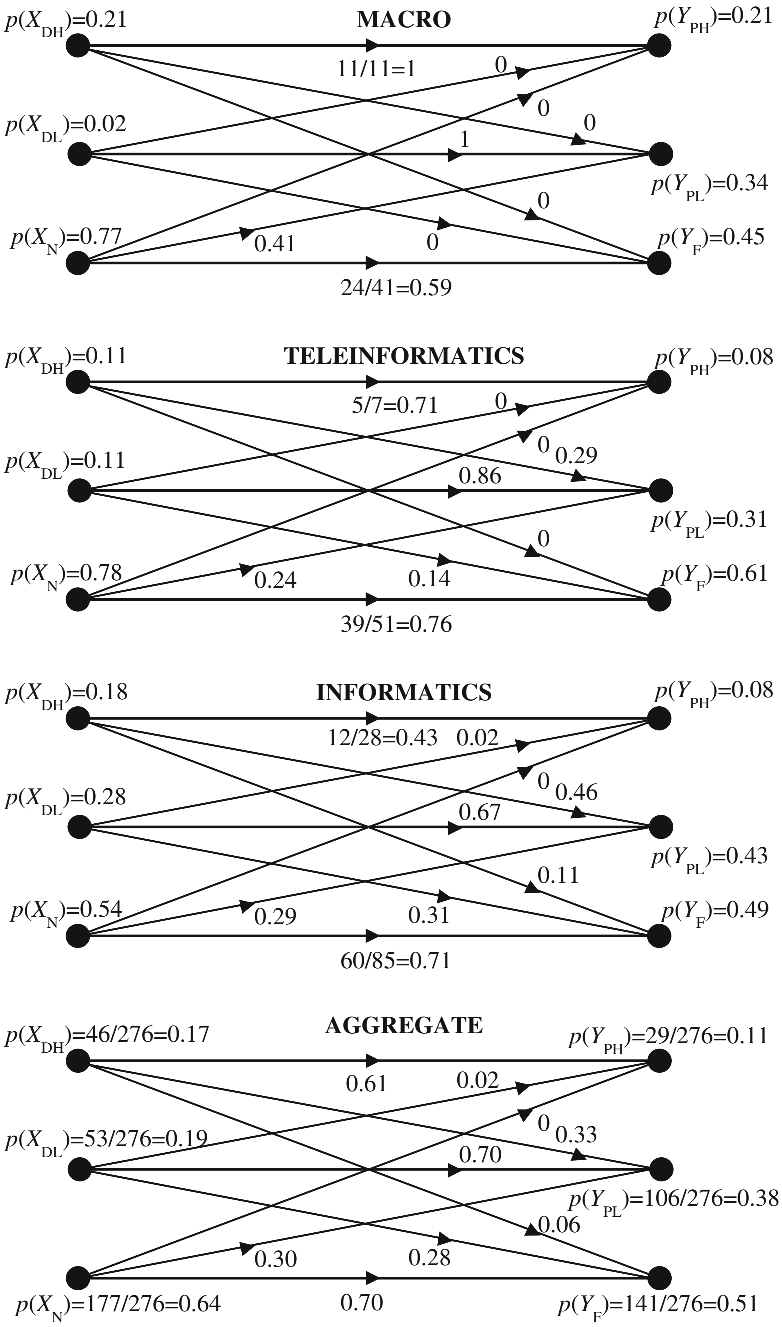

Fig. 4. Case study 2015/2016 3-input/3-output channels for 3 fields of study and aggregate

Table 2. Conditional probabilities of 3-input/3-output aggregate Case Study Channel

$p(Y_j/X_i)$	$Y_1 = Y_{PH}$	$Y_2 = Y_{PL}$	$Y_3 = Y_F$
$X_1 = X_{DH}$	28/46 = 0.61	15/46 = 0.33	3/46 = 0.06
$X_2 = X_{DL}$	1/53 = 0.02	37/53 = 0.70	15/53 = 0.28
$X_3 = X_N$	0/177 = 0	54/177 = 0.30	123/177 = 0.70

4 Final Conclusions and Guidelines

The updated new method that enables quantitative evaluation of effectiveness of Formative tests (quizzes) has been proposed. All three findings of the 2014/2015 case study have been confirmed. Taking also into account 2015/2016 case study, the following final findings can be formulated:

1. Number of Negligent students, students that started Formative quizzes just (one week) before the deadline and solved* only the obligatory minimum of tasks (20%) is high, ranges from 54% to 78%, with 64% (≈2/3) as the average!!!
2. Number of students that started Formative quizzes just before the deadline is even greater, and reaches 75% (11% managed to complete T_{DL} = 33% of tasks during the last week before the deadline to become Diligent Low).
3. None of Negligent students have Passed with High mark.
4. 70% of Negligent students Failed**.
5. Only 6% of Diligent High didn't Pass***.
6. 61% of Diligent High Passed with High mark.
7. Information channels for different fields of study are very similar.

* It can be suspected that some Negligent students cheated when passing obligatory Formative quizzes, used cheat-sheets (repository of answers-formulas) or asked peers to do the job for them.
** Significant percentage (30%) of Negligent students Passed Low only due to low Pass threshold of 3.5 points, for the threshold set on 5 points (50%) only 5% of Negligent students would Pass.
*** It has been observed that some students classified as Diligent High have passed Formative quizzes without understanding, just using cheat-sheets. Then, they have learnt answers-formulas by heart hoping that it will be enough to set the Exam-data to these formulas and Pass without understanding. Unfortunately for them, when preparing the Exam quiz, some minor corrections have been introduced to Formative questions such that the formula learnt by heart didn't give the correct answer.

These findings gave valuable feedback to the teacher:

- Proved high quality of the designed Formative online quizzes, usefulness of e-materials supporting SDL in the Flipped Classroom (video-podcasts, screencasts, e-textbook, e-slides explaining reasoning).
- Proved proper alignment of the Formative online quizzes and the final Exam and consequently its compliance with the teaching goals and learning outcomes.

- Made it clear that there is an urgent need for greater motivation of students to systematic work and solving quizzes with understanding.

To meet this need, some necessary steps have to be undertaken. In the carrot and stick strategy applied more stick has to be added, and the following final and unconfirmed guidelines can be formulated. To improve Pass/Fail ratio, not violating achievements of learning outcomes:

- It is necessary to force students to more systematic work. To reach this goal:
  - All Formative quizzes have to be obligatory, i.e. distribution of questions answered has to be uniform in the set of all quizzes.
  - Common deadline on all quizzes at the end of semester has to be replaced by deadlines on individual quizzes, distributed uniformly during the whole semester, as the lecture goes on.
  - Systematic work, solving quizzes with understanding, has to be verified during classroom tutorials.
  - Top students, some 10% of students that solved systematically the greatest number of tasks may obtain upgrade of the Exam grade or even get the credit, with Excellent grade, based on Formative quizzes alone.
- It is necessary to persuade students that only solving problems with understanding has sense. Completing quizzes using cheat-sheets (repository of formulas) is self-deception and also rote learning of these formulas will not pay. The Exam questions have to differ slightly from Formative questions, such that stored (learnt by heart) formulas are useless.
- The opportunity for cheating, both while solving Formative quizzes and solving the Exam quiz has to be reduced to minimum. What regards Formative quizzes, solving with understanding can be verified during classes, e.g. by short, single question tests. What regards the Exam, some additional technical precautions in a computer lab have to be undertaken, first of all to prevent from using mobile devices.

References

1. Rutkowski, J.: Flip teaching approach in smart education - principles and guidelines. Keynote of the 2nd International KES Conference on Smart Education and E-learning, Sorrento (2015). https://drive.google.com/file/d/0B_esc-ATuPmhX05RXzBKRzZIVTg/view
2. Fielding, A., Bingham, E.: Tools for computer-aided assessment. Learn. Teach. Action **2**(1) (2003). http://www.celt.mmu.ac.uk/ltia/issue4/fieldingbingham.shtml. Learning and Teaching Unit of Manchester Metropolitan University
3. Sim, G., Holifield, P., Brown, M.: Implementation of computer assisted assessment: lessons from the literature. ALT-J Res. Learn. Technol. **12**(3), 215–229 (2004)
4. Computer Assisted Assessment – Research into E-Assessment. In: Proceedings of the International Conference on CAA 2014. Communications in Computer and Information Science Series, Zeist. Springer (2014)
5. Rutkowski, J.: Moodle-based computer-assisted assessment in flipped classroom. In: Smart Education and Smart e-Learning. Smart Innovation, Systems and Technologies, vol. 41, pp. 37–46 (2015)

6. Uskov, V., Bakken, J., Pandley, A.: The ontology of next generation smart classrooms. In: Smart Education and Smart e-Learning. Smart Innovation, Systems and Technologies, vol. 41, pp. 3–14 (2015)
7. Nicol, D.J., MacFarlane-Dick, D.: Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. Stud. High. Educ. **31**(2), 199–218 (2006)
8. Rutkowski, J.: Evaluation of the correlation between formative tests and final exam results – theory of information approach. Int. J. Electron. Telecommun. **62**(1), 55–60 (2016)
9. Guglielmino, L.M.: Reactions to field's investigation into the SDLRS. Adult Educ. Q. **39**(4), 235–240 (1989)
10. Higgins, M., et al.: Effective and Efficient Methods of Formative Assessment, CEBE Innovative Project in Learning & Teaching (2010). www.smu.ca/webfiles/Effective_and_Efficient_Methods_of_Formative_Assessment.pdf
11. Rutkowski, J.: Theory of Information and Coding. The Publishing House of The Silesian University of Technology, Gliwice (2006)
12. Moodle 2.7 Documentation. https://docs.moodle.org/27/en/Main_page
13. Rutkowski, J., Moscinska, K., Chruszczyk, L.: Development and evaluation of computer assisted exam – circuit theory example. In: Proceedings of the 9th IASTED International Conference on CATE, Lima, Peru, pp. 333–338 (2006)

http://www.springer.com/978-3-319-59450-7

Smart Education and e-Learning 2017
Uskov, V.L.; Howlett, R.J.; Jain, L.C. (Eds.)
2018, XIII, 498 p. 186 illus., Hardcover
ISBN: 978-3-319-59450-7