Software Solution Improving Productivity and Quality for Big Volume Students' Group Assessment Process

Martin Misut
University of Economics in Bratislava, Bratislava, Slovakia
martin.misut@euba.sk

Maria Misutova
Slovak University of Technology in Bratislava, Trnava, Slovakia
maria.misutova@stuba.sk

Abstract—The key difference between testing a small and a large group of students is typically the impossibility to use an e-assessment interface for a large group of students in one room, especially in mathematics courses. Traditional exam with written tests causes a burden on the teacher when evaluating students´ results, and a delay in releasing the results. Moreover, the detailed results and test content are not directly stored in the database for further evaluation, which is a condition for optimization of the learning environment. A new teaching model and a method of assessment with enhanced software support were designed to overcome these problems. Experimental data collected over three experimental academic years were compared to traditional assessment data to evaluate the efficiency of the proposed solutions. Results have proved that the overall productivity was significantly increased – more than eleven times without compromising quality. Internal dependencies among parameters of the automated assessment procedure were identified, analyzed and used to maximize the impact of automation, as well.

Keywords— assessment, assessment process, automation, mathematics, productivity

1   Introduction

Education has been increasingly relying on the use of technology. In the past decade, educators and administrators have used Student Information Systems (SISs), Learning Management Systems (LMSs), and enrollment and admissions management systems to optimize and streamline classroom, school, and campus operations. [1] As it was stated in Big Data in Education report [1], recently mobile devices, online assessments, digital curricula, education applications and digital whiteboards had been adopted worldwide by educational institutions to advance teaching as well as learning.

Summative assessment, which is one of the objects of technology usage in education, has had a special place within digitalization of education [2]. Assessment procedures are being adapted to new learning environments by moving from isolated summative assessments to more integrated assessment forms, illustrated in some studies, for example in [3], [4] [5], [6], [7], [8]. The term e-assessment is often used to emphasize novelty and on-line technology utilized in the assessment. E-assessment can be defined according to [9] as a process where information and communication technology is used for the management of the end-to-end assessment process. In other words, e-assessment deals with methods, processes and web-based software tools (or systems) which allow systematic inferences and judgments to be made about the student's skills, knowledge, and capabilities.

E-assessment includes two different settings as stated by authors in [7]: the physical classroom (an assessment with the support of technology) and online classroom (assessment through technology). Furthermore, there are also two types of classrooms in the online environment setting: the asynchronous on-line classroom and synchronous on-line classroom. Therefore, it can be concluded that there are three different possible settings of the e-assessment in the modern ICT era: assessment in the traditional physical classroom, in the asynchronous on-line classroom and synchronous on-line classroom [7].

E-assessment offers several advantages in comparison to the traditional paper-based exams: the grades and feedback can be provided immediately, and there is greater flexibility in choosing the location and time of the exam. Other advantages of e-assessment, such as presented by Fuentes in [3], include time and cost effectiveness, the fact that interactive activities and multimedia tools can be incorporated into the assessment process, and that teacher's interpretations and legibility problems are avoided among others. Moreover, as Daly et al. [10] indicate, the use of ICT in assessment blurs out the boundaries between formative and summative evaluation. Several studies have provided evidence (e.g. [11]) that online assessment methods can provide a valid measure of performance, equivalent to that obtained using traditional methods. Such evidence is relevant in motivating and justifying the use of e-assessment in line with [12], especially in summative contexts. Students' experience of summative assessment has been the focus of many important and rightly influential studies that confirm the central role played by assessment in students' academic lives [12].

One of the e-assessment's significant advantages is that it allows personalized testing where the choice of the learning object or test item is adapted to the needs or knowledge of the learner [5]. The adaptive e-assessment can be used to assess specific knowledge and competences that the learner has acquired up to the current point of the learning process [9]. Because of the rapid development of the formative use of e-assessment, the adaptability is becoming an essential feature of e-assessment. The flexibility of e-assessment increases because students can use the feedback to adapt their approach to performing tasks. The key condition for adaptability of the assessment is an automated scoring ability of the e-assessment system [9].

Mathematics, including basic statistics, is a particularly important subject both in primary and higher education. As mathematics forms a fundamental component of all

science, technology and engineering disciplines; mathematics courses are often a compulsory part of study programs at universities, especially in the engineering education. Numerous research studies have shown that the use of instructional technologies helps to improve the teaching and learning processes of mathematics [13]. Webb [14] showed that assessment in math courses is fundamental to the whole educational process. Therefore, there has been an extensive effort in the recent years to identify methods and possibilities of the instructional technologies implementation to support assessment as an integral part of the learning process (e.g. [6, 10, 15-17]).

As Sangwin and Jones declared in [18], in mathematics, assessments typically attempt to measure one or both the procedural knowledge and the conceptual understanding. Haapasalo in [19] defines procedural knowledge in mathematics as knowledge that denotes dynamic and successful use of specific rules, algorithms or procedures within relevant representational forms. The same author explained conceptual knowledge as knowledge that denotes knowledge of particular networks and a skillful "drive" along them. One of the reasons for emphasizing procedural items when appraising the assessment of mathematics is that they are relatively easy to produce and can be scored objectively [20]. Another reason for the emphasis on procedural items is that the technical and engineering education that embraces an engineering-oriented perspective as an avenue for the development of the meaningful and authentic problem-solving capabilities in students lies mainly on procedural knowledge of mathematics.

The extended use of technology results in a frequent usage of tests for assessment as well [21]. Two most frequently used common test question formats for procedural test items in mathematics are multiple choice questions (students have to choose the correct answer from a number of options) and constructed response. Procedural multiple choice items typically present a mathematical object, such as an equation, and an instruction to transform the object into a specified form [18]. The equivalent constructed response item, as Sangwin and Jones stated in [18], would contain the same question item, but the answer options would be removed and replaced with either a space to write down the answer or with a text box in the case of computer-based assessment.

Some early assessment systems relied only on multiple choice and numeric input questions. Sangwin and Köcher in [22] pointed out that even these simple forms of e-assessment have been useful in different aspects of the assessment of mathematics. The interface of multiple choice and numeric input assessment systems is quite easy to use even for users with only basic computer experience. In constructed response items systems, students can enter the answer to a particular question and the system will then evaluate it [23]. The constructed response items raise the difficulty of students needing to learn specific syntax to enter their answers. The need to learn the new syntax could hinder the performance of all or some of the students [18]. Moreover, the automatic scoring function of the e-assessment system is much harder to implement for constructed response than for multiple-choice or the numeric input items.

The evaluation of assignments and tests scoring is only one part of the whole assessment process as other time-consuming tasks should be carried out [24].

Preparation of knowledge base that can be used to assess the students' outcomes properly is a tough and time-consuming task that requires an extensive amount of resources [25]. E-assessment with automated scoring, in comparison to traditional assessment with written tests, enables to move the teacher's attention to assessment preparation and the evaluation of results. The traditional assessment, on the other hand, causes overloading of the teacher when evaluating students´ results and a delay in releasement of the results [26]. Furthermore, the absence of data about student answers in digital form makes the use of test results and the optimization of tests almost impossible. This applies particularly to the assessment of a large group of students[1] when assessing them several times during the semester. Furthermore, in most cases, the final summative e-assessment realized for students´ grading purposes is organized for students in groups under the inspection of the teacher to prevent cheating. Under the normal conditions, there are not any problems with the exam assessment since the number of students is small, probably up to thirty students.

Every academic year, several hundreds of students attend in the first two semesters the core engineering courses such as mathematics and physics at the Faculty of Material Science and Technology in Trnava. Because of a vast number of students, the examination during the semester as well as the final exam cannot be organized in small groups, but only for groups of often more than two hundred examinees. The number of students implies that there are unique technological (and other) requirements that should be fulfilled when organizing exam since the conditions are quite different from the usual settings when working with groups of less than thirty students. One of the fundamental differences between an exam with a considerable number of students and a small number of students is the impossibility to use an online assessment directly for a large group of students. Thus the difficulty with the automated tests scoring as well as the evaluation of the results arises.

The philosophy of the solution that was proposed and implemented by the project research team will be described in the next two chapters. Three years of experiences of the proposed software assessment together with a discussion of the usefulness of proposed solution will be presented in the rest of the paper.

2 Problem definition

The teaching of mathematics in engineering studies has been worldwide reported as problematic due to the low level of student retention rate, e.g. [27] or [28]. The same situation is experienced at the Faculty of Material Science and Technology as published in [29] or [30]. Different tools and steps have been implemented at universities throughout the world to maximize student retention rate. Some of the possibilities can be found in e.g. [31], [32], [33], [6], [34], [35], [36], [37]. Proposed solution usually includes merging technology and the results of pedagogical research which, in turn, results in intensive student support, formative e-assessment, and

---

[1] more than four hundred

encouragement to the students' use of their thinking and thus personalization of the learning environment.

To improve retention rate and to support and enhance students´ math skills development at the universities, we have designed a new teaching model. [38] The new design puts emphasis on continual learning, activity, independence, and creativity [39] and contains education and assessment blocks and a database of tasks that require the application of knowledge. Technology has supported the educational process to a great extent. The developed teaching model was applied in the mathematics courses during the 2012/2013 academic year in the first year of the study. Based on the obtained results, the model was adjusted, and the new version was implemented from the winter semester of the 2013/2014 academic year. The assessment which was a part of the new teaching model was substantially redesigned and adapted to the use of technology. However, two design problems have occurred: a problem with assuring the automatic scoring in the summative assessment and a problem of choosing the interface of the assessment system that would enable assessment of several hundreds of examinees during the exam under an inspection of the teacher to prevent cheating.

3 Software solution architecture

The main idea of the presented solution is based on the digitalization of answer sheets that are produced as a result of the offline tests. These offline tests are administered, prepared and evaluated by the software. It means that the transformation between the digital and non-digital form of the test and vice versa occurs two times in one cycle, as it can be seen in Fig.1.

The philosophy of assessment automation for significant size student groups covers several activities and artifacts. Firstly, the lecturer[2] prepares a database of questions that can be used in exams (Activity A in Fig. 1). The preparation of the database is a tough and a time-consuming task. [25] Afterwards, the structure of the test is defined and stored in the database (Activity B). The prepared test structure is used for the generation of different versions of tests for each student in the next step (Activity C). Each copy of the test is stored in the database, waiting for answers from students. All these activities should be realized before the exam day. The printing of test sheets (Activity D) has to be finished before the exam will start. Signs for digital identification of test and student are added to each answer sheet. Students write down their name, ID and test code. Afterward, the student´s answer sheet is paired with a particular version of the test and student´s record by their ID and test code. Students write their answers on the answer sheet (Activity E), and hand the answer sheet to the teacher at the end of the exam. Answer sheets are then scanned (Activity F) to digitalize the answers and afterward sent in batches to the AIS server. (AIS is the abbreviation of the name of student information system – Academic Information System.) Scanned batches of answers are analyzed with the purpose of identifying the

---

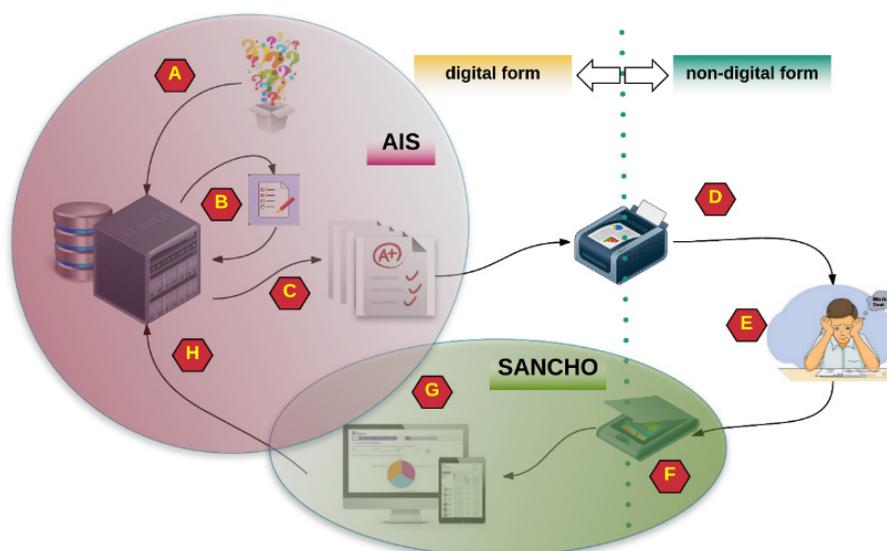[2] Terms lecturer, examiner and assessor are used as synonyms of teacher

Fig. 1. Philosophy of off-line assessment for big volume students´ groups

answers and assigning a particular version of the test to student´s record on the base of student's ID and a test code. If errors occur, they are manually corrected (Activity G). Results of the student's answer identification and student assignment are stored for each student in the database, similarly as the digital test, for later evaluation and student grading (Activity H). Exam results are available for follow-up analysis as well.

Technically, the software solution is in the form of a client-server application named SANCHO that extends the functionality of AIS, the information system already used by the university. SANCHO was developed by IS4U Company as an add-on module of the Academic Information System to support scanning of the written exams.

Company IS4U was selected because it was necessary to implement new functionality in AIS and this company is the vendor of AIS. The new software application was developed according to our requirements. The SANCHO application includes tools that, combined with the input device (scanner), allow a very easy transfer of written tests into the digital form that is required for automated processing. Since SANCHO and the university's information system (AIS) are closely related, it offers the most efficient and comprehensive tool to users for creating, managing, scanning and evaluating all kinds of tests, which are currently used at colleges and universities.

Client-server communication is assured through web services. Those web services ensure the creation, transfer, and deletion of batches of scanned tests, authentication of users, managing the transfer of batches to AIS as well as other operations with batches in AIS such as batch closure, its´ handing over to the teacher or pairing with courses.

The client application is the MS Windows type of software with the GUI providing the acquirement of the digital form of tests, the administration of the tests and transfer of the tests to the server. The server then processes and evaluates the scanned tests. It is possible to present the results of scanned tests over the unified user interface of AIS through the direct integration with AIS, which students and teachers are already familiar with.

The client layer of the application is written in C # on Microsoft .NET 4.0 framework. MS Windows XP SP3 with the framework of Microsoft .NET version 4.0 is a minimum requirement. Thanks to the use of .NET framework, SANCHO has a similar user interface to what users already know from some Microsoft products.

The user might not have access to a local network or the Internet when scanning the documents and creating batches of scanned tests. The entire scanning process can be performed without network access. However, an actual transfer of batches to the server requires network connectivity. On the other hand, if the user is connected to the network and authenticated in the AIS during scanning, scanned documents are automatically synchronized with the server. Further processing then takes place on the server. The SANCHO client primarily serves for scanning and managing scanned documents and their batches.

However, scanning is only part of the whole process followed by other operations especially knowledge mining, data processing of scanned papers and tests evaluation. Thanks to the support of TWAIN and ISIS interface, SANCHO allows any scanner to be used for the digitalization of documents. Nevertheless, for greater convenience, the recommended devices should be equipped with an automatic feeder. [40]

The server layer of the application is implemented in Perl as an extension module of the Students' Information System (AIS). Communication with the client takes place via web service that provides the takeover of sent batches of scanned documents and other minor operations such as creating and removing batches. The advantage of the integration of the server into the information system is the ability to store received batches directly into the database of the student information system, where the batches are ready for further processing and use in SANCHO or the information system.

The primary task of the server layer is the analysis of scanned documents, the identification of individual parts that each answer sheet is consisted of, the recognition of questions and replies to these questions, and lastly, the transfer of answers from scanned documents into an electronic form that will be stored in the information system. Each batch will be assigned to the appropriate exam and analysis process will then start. The following sequence of operations are applied to each sheet in the batch:

1. Normalization - automatic optimization of brightness and contrast to facilitate subsequent analysis of the document.
2. Identification of positioning signs - the test sheets are provided with a set of positional markers that enable the determination of the basic geometry of the document. They serve for performing the correction of skew, rotation, zoom in or out and the sheet orientation.

3. Identification of the document - once the sheet is positioned, it is possible to start the identification process. This process is based on the recognition of barcodes that are printed on each sheet. The bar code is used for the detection of the test, type of the sheet and detection of a page number.

4. Recognition of the document parts - a sheet type activates the detection of document elements. Identification of individual questions and their corresponding answers is done directly for separate answer sheets. For sheets containing an answer field directly nearby a question, localization of questions based on barcodes is applied first, and then the answers are recognized. Direct analysis of student's responses and transfer of the responses into the digital version of the test is done for the simple choice type of questions. This kind of question can be evaluated automatically. Open-ended questions that require evaluation by the teacher are included into the digital form of the test as a picture.

5. Evaluation – the assessment tool evaluates recognized answers stored in the digital version of the tests. For questions for which the answers have failed to be correctly assessed, the viewport of answer sheet with the respondent's answers is stored as a picture. In case the answer is not recognized, the examiner is notified, and that particular area of scanned document is displayed when checking the test results. Therefore, the examiner has the possibility to evaluate the answer correctness manually using the tool for evaluation of tests.

The evaluation finishes, once all the questions of the test have been processed. Further processing depends on the examiner's needs. Tests in digital form are possible to be viewed and reviewed with the built-in tool. It is possible to create output reports in MS Excel format for future processing (widely used in surveys and polls) or to generate reports in PDF data file type which is particularly suitable for the publication of results. One of the advantages of the presented solution is the integration of all software apps such as individual modules of SANCHO, AIS, e-learning tool, and student assessment instruments. This integration allows the examiner to transfer the test results directly into interim evaluation or to the overall assessment of students in the course. Basic statistical data (such as percentage of successfully passed tests, average, median, minimum and maximum points achieved in the test, details of the frequency of question usage for each category, number of questions used, answer correctness ratio for each question, and answer selection frequency) is available for each test. The data can be processed further, can be filtered, sorted by various user defined criteria, exported, and visualized into a graph, and so on.

4 Design of the Experiment

The regular use of new software started in the academic year 2013/2014 after the implementation of SANCHO, and when user tests and user training had finished. The experiment was designed to evaluate the usefulness and effectiveness of the newly implemented software solution. The research question has been formulated as follows:

Is the newly implemented procedure of massive off-line student assessment tool in the course Mathematics I more efficient than the traditional method of assessment? If yes, to what extent?

The effectiveness of the assessment method here is considered with the focus on automation of the procedure itself. The effectiveness and validity of tests for the course Mathematics I were examined as well and are published separately, e.g. in [29].

The effectivity of the chosen assessment method needed to be defined to answer the research question. Assessment method effectivity can be easily measured by the number of tests completely processed over one-time unit (productivity) in compliance with quality at 100%. The criterion of quality for the purpose of research experiment was defined that all of the answers of all of the tests should be truly recognized, fairly evaluated and correctly stored in AIS for each student tested. Various kind of errors have appeared during the recognition of answers from the answer sheets, so the time to correct these errors has been taken into consideration when calculating the productivity of the assessment method.

The research was conducted over three consecutive academic years 2013/14, 2014/15 and 2015/16. The studied sample consisted of students in the course Mathematics I during all three years of experiment. Stored data from 2012/13 academic year has served as a control sample for the experiment.

From the academic year of 2012/13, the assessment of the course Mathematics I has consisted of several interim tests and a final exam test. In case that a student does not pass their final exam test, they have the possibility to repeat it two times. Therefore, the total number of final exam attempts is higher than the number of tested students.

Traditional assessment of the course Mathematics I was in the form of math problem solving for the written exam. The exam questions/problems have always been newly prepared for every exam, and afterward, the exam sheets were printed out for each student. Prepared exam sheets were distributed right before the exam started. The duration of all of these activities is presented in Table 1 in the column with the preparation time heading. Answer sheets were collected after the exam finished. In the next step, the answer sheets were evaluated by seven teachers and the grades of all students that haven taken the exam have been manually inserted into the AIS database. All post-exam activities are considered here as processing activities.

The experimental preparation procedure of exam sheets can be seen in Fig. 1 (Activities A to D). Moreover, the distribution time of both the printed exam sheets and the answer sheets was added to the length of activities A to D to calculate the preparation time for the corresponding years (as seen in Table 1). The sum of the duration times of activities F to G (Fig. 1) is denoted as assessment processing time for the particular academic year.

The content of the course Mathematics I has been unchanged for all academic years considered and covers following areas: functions of a real variable, functions of several real variables, determinants, systems of linear equations, complex numbers, and optimization of functions of a real variable.

Tests containing different types of tasks with a merged answer sheet were used in the academic year 2013/14. These tests were printed on three pages of two paper sheets. Previous experience with this kind of tests showed us the necessity to scan two sheets in the scanner duplex mode per student. To lower the scanning time, another type of test was used from the academic year 2014/15, where the universal answer form is printed on one sheet of the paper (see Table 3).

Examining large group of students, (more than 200) required careful preparation and excellent organizational skills. The preparation of the exam in Mathematics I included the development of the seating plan that was sent to students one day before testing, the preparation of the content of the exam, and the printing tests and answer sheets that were distributed in the examination room according to the seating plan.

5   Obtained results

The cumulative duration of times of the pre-test activities, as well as post-test activities for all test runs in given academic year, are presented in Table 1. The length of all activities was measured in minutes and then converted into hours.

The data in Table 1 should be considered together with data from other tables to understand the nature of the data in Table 1 fully. The proportion of processing time (post-test activities) has been dramatically reduced from 93.0% to 45.6% together with a significant decrease of the duration of the post-exam activities as a result of the assessment automation. Reducing the proportion of processing time means that the exam results are ready sooner and that lecturer's workload is more balanced compared to the situation before Sancho utilization (7:93 –> 54,4:45,6). The obtained data supports the teacher´s observation from experiments as well. The data indicates that the teacher's workload has been moved from processing of the exams to the exam preparation and at the same time both portions of the assessment (preparation and processing) have been shortened.

Unitary assessment productivity (UAP) was defined as the number of completely prepared and assessed student exam attempts completed in one hour to measure the assessment productivity. UAP is calculated as the total number of student exam attempts in one academic year divided by total time of all activities except for the exam activity time for all student exam attempts in one academic year.

Values of productivity for three experimental academic years together with a control year data for the traditional assessment method are presented in Table 2. As it is evident from the traditional assessment method data, the preparation productivity was slightly improved thanks to the software support of test generation. On the other hand, the productivity of post-test activities was increased by more than 23 times (see tab.2) and overall unitary productivity has improved by more than 11 times (from 4,9 up to 56,6). The overall assessment productivity of seven teachers in the academic year 2012/13 was lower (34,5) than the assessment productivity of only a single teacher using the software tool in the academic years of 2013/14 and onwards.

The following three types of errors were recognized during the academic year of 2013/14 (the first year of the new system operation):

1. Erroneously recognized student identification number (ID), i.e. the software was not able to assign a student name to a given test according to prescribed ID. This type of error has been manually corrected by assigning the student name to an answer sheet by selecting the student from a student list. Since the manual selection was a time-consuming activity, a SANCHO modification was proposed which led to a significantly accelerated correction of this error type. The proposal was accepted and implemented into SANCHO immediately.
2. Incorrectly recognized data, i.e. some answers were not identified in the answer sheets. A red question mark indicated missing data in the list of results. Fields with unrecognized answers are color-coded. A correction was made manually by the lecturer who would check and make corrections of missing answers in the scanned document that has been attached to the digital version of the test. Subsequently, the test is designated as reviewed.
3. Scanning errors, i.e. the content of a scanned sheet was missing. A red letter x has indicated the missing content in the list of results. This type of errors has only occurred in the first period of testing, ergo in the academic year of 2013/14, and afterward, a software update removed the errors.

The frequency of different types of errors in cumulative and relative forms are presented in Table 3. It can be seen from the Table 3 that the cumulative frequency of errors was lowered nearly by half (from 30.69 to 17.08 per hundred of student exam attempts) due to the increase in quality of scanning. The number of errors directly influences the manual part of a teacher´s work during the assessment, and therefore, the workload of the assessor will be influenced as well. The reduction of all error types belongs to tasks with the highest priority in the automation of assessment.

Table 1. Time structure and characteristics of assessment in different academic years

Ac. Year	Number of test runs	Number of student attempts	Preparation time [h]	Processing time [h]	Total time [h]	Proportion of processing time [%]
12/13	6	543	7,7	102,4	110,0	93,0%
13/14	5	404	4,6	5,7	10,3	55,3%
14/15	9	1271	12,2	10,8	23,0	46,9%
15/16	9	1259	12,1	10,1	22,2	45,6%

Table 2. Assessment productivity in different academic years

Ac. Year	Unitary productivity of preparation [test/h]	Unitary productivity of processing [test/h]	Unitary assessment productivity [test/h]	Number of teachers	Overall productivity of processing [test/h]	Overall assessment productivity [test/h]
2012/13	70,8	5,3	4,9	7	37,1	34,5
2013/14	88,1	71,2	39,4	1	71,2	39,4
2014/15	104,4	117,9	55,4	1	117,9	55,4
2015/16	104,1	124,1	56,6	1	124,1	56,6

The initial recommendation for scanning resolution was 200x200 dpi, and this resolution was used in the academic year of 2013/2014. However, due to the high frequency of all three types of errors, the resolution has been improved in subsequent years to 300x300 dpi and then to 400x400 dpi. The corresponding data is shown in Table 4.

Table 3. Statistics of post-test processing errors

Ac. Year	Number of student exam attempts	Number of sheets	ID errors		Data errors		Scanning errors		All errors	
			total	*per 100 attempts*	*total*	*per 100 attempts*	*total*	*per 100 attempts*	*total*	*per 100 attempts*
2013/14	404	808	75	18,56	25	6,19	24	5,94	124	30,69
2014/15	1271	1271	184	14,48	152	11,96	0	0,0	336	26,44
2015/16	1259	1259	170	13,50	45	3,57	0	0,0	215	17,08

Table 4. Dependence of processing errors on the scanning quality

Academic Year	Scanning quality (dpi)	Scanner speed (sheets/min)	Processing time total [h]	Manual corrections [h]	Manual corrections (%)
013/14	200	27	5,67	3,48	61,29%
2014/15	300	12	10,78	5,09	47,20%
2015/16	400	7	10,14	2,95	29,08%

6   Discussion

Both the analysis of the assessment activities as well as the obtained experience from three years of SANCHO usage have resulted in the conclusion that the parameter that can influence the productivity the most is the quality of scanning, expressed in dpi (see the tab. 4). The quality of scanning will directly influence the scanning speed (tab. 4), and also will affect the activity of recognition indirectly through the frequency of errors (see tab.3) depending on the scanning quality. The quality of scanning and recognition will influence assessment productivity. Further increase of quality of scanning (500x500 dpi) has resulted in a considerably lower speed (only three sheets per minute) therefore, this resolution was excluded from the experiment. To further reduce the number of ID errors, a new method of exam preparation should be adopted. Such method should consist of generation, printing, and delivery of personal answer sheets before the examination. This type of arrangement would remove the ID errors because bar codes would make the sheet identification. On the other hand, the proposed method would probably consume more teacher time before the examination to ensure that each student has been assigned their personal answer sheet.

Further elimination of data errors could be done by a more consistent filling of the answer sheets. Despite being instructed not to, the students have written some letters or numbers by writing the line several times when marking the answers, which

resulted in data errors during the recognition activity. Since further reduction of data errors is probably not possible to reach by improving the student's writing, the only possibility that is left is to rewrite the recognition algorithm which is a tough task. As Suzuki et. al. [41] pointed out, although many OCR technologies for math content have been studied for a long time, only the InftyReader (the OCR system for math documents) has been put to practical use. InftyReader could properly recognize only documents in print or PDF, including math expressions, tables, graphs and other technical notations if their layout is not very complicated. The real complication with recognition of math expressions means that it is better to keep the test format when assessing a significant group of students.

The decrease in the time of human intervention into recognition activity has a higher value for a teacher than the reduction of total time of processing. When comparing the 2014/15 and 2015/16 academic years, it is evident that the total processing time is approximately the same but the time of manual corrections has dropped by more than two hours (tab. 4). This behavior is desirable because the teacher can do something else during the activity of automatic scanning and processing of the answer sheets.

7 Conclusion

The main benefit of assessment automation lies not only in the time reduction during the test scoring but also in a significant decrease of the likelihood of errors committed by the student due to an incorrect writing of numbers and letters. Currently, a large number of students have illegible handwriting as a result of excessive usage of IT devices.

Finally, it can be stated that the proposed, tested and used system of assessment for the continuous and final exam testing in the course Mathematics I is suitable in this form. Based on the obtained experience, the assessment system can be recommended for use and evaluation of tests as well as for verification of knowledge in other courses attended by a big number of students. At the same time, it should be noted that the usage of automated test assessment enables teachers to concentrate on the exam preparation more precisely and thus lowers their post exam stress.

Currently, the answers to assignments and exams are the only measurements of student performance. Nevertheless, every student will generate a unique data trail during their student life. This data track can be analyzed in real-time to deliver an optimal learning environment for the student as well as to gain a better understanding of the individual behavior of students. The first step of analysis has already been done by the implementation of the proposed solution as described in this paper. We believe that by delivering an optimal learning environment to the students, it will also result in an increase of the retention ratio of first-year students.

8   Acknowledgment

9   References

[1] Reid-Martinez K., Mathews M. (2015). Big data in education: Harnessing Data for Better Educational Outcomes. Center for Digital Education, 100 Blue Ravine Road, Folsom, CA 95630. pp. 33.

[2] Webb M., Gibson D., Forkosh-Baruch A. (2013). Challenges for information technology supporting educational assessment. Journal of Computer Assisted Learning, 29:451-462. https://doi.org/10.1111/jcal.12033

[3] Fuentes J.M., Garcia A.I., Ramirez-Gomez A., Ayuga F. (2014). Computer-based tools for the assessment of learning processes in higher education: A comparative analysis. INTED2014: 8th International Technology, Education, and Development Conference, March 10-12 2014, Valencia, Spain. pp. 976-984.

[4] Gusev M., Ristov S., Arminski G., Velkoski G., Bozinoski K. (2013). e-Assessment Cloud Solution: Architecture, Organization, and Cost Model. International Journal of Emerging Technologies in Learning (iJET), 8:55. https://doi.org/10.3991/ijet.v8iS2.2783

[5] Wesiak G., Al-Smadi M., Höfler M., Gütl C. (2013). Assessment for complex learning resources: Development and validation of an integrated model. International Journal of Emerging Technologies in Learning, 8:52-61. https://doi.org/10.3991/ijet.v8is1.2354

[6] Diefes-Dux H.A., Zawojewski J.S., Hjalmarson M.A., Cardella M.E. (2012). A Framework for Analyzing Feedback in a Formative Assessment System for Mathematical Modeling Problems. Journal of Engineering Education, 101:375-406. https://doi.org/10.1002/j.2168-9830.2012.tb00054.x

[7] Chao K.J., Hung I.C., Chen N.S. (2012). On the design of online synchronous assessments in a synchronous cyber classroom. Journal of Computer Assisted Learning, 28:379-395. https://doi.org/10.1111/j.1365-2729.2011.00463.x

[8] Hostovecky M., Misut M., Pribilova K. (2015). Web-based testing in science education. Lecture Notes in Electrical Engineering, 313:247-254. https://doi.org/10.1007/978-3-319-06773-5_33

[9] Baneres D., Baró X., Guerrero-Roldán A.E., Rodríguez M.E. (2016). Adaptive e-assessment system: A general approach. International Journal of Emerging Technologies in Learning (iJET), 11:16-23. https://doi.org/10.3991/ijet.v11i07.5888

[10] Daly C., Pachler N., Mor Y., Mellar H. (2010). Exploring formative e-assessment: using case stories and design patterns. Assessment & Evaluation in Higher Education, 35:619-636. https://doi.org/10.1080/02602931003650052

[11] Hewson C. (2012). Can online course-based assessment methods be fair and equitable? Relationships between students' preferences and performance within online and offline assessments. Journal of Computer Assisted Learning, 28:488-498. https://doi.org/10.1111/j.1365-2729.2011.00473.x

[12] Joughin G. (2010). The hidden curriculum revisited: a critical review of research into the influence of summative assessment on learning. Assessment & Evaluation in Higher Education, 35:335-345. https://doi.org/10.1080/02602930903221493

[13] Andrade-Aréchiga M., López G., López-Morteo G. (2012). Assessing effectiveness of learning units under the teaching unit model in an undergraduate mathematics course. Computers & Education, 59:594-606. https://doi.org/10.1016/j.compedu.2012.03.010

[14] Webb N.L. (1992). Assessment of students´ knowledge of mathematics: steps toward a theory. In: D. A. Grouws (Ed.), Handbook of research on mathematics teaching and learning, NTCM/Macmillan, New York. pp. 661-683.

[15] Blanco M., Ginovart M. (2012). On How Moodle Quizzes Can Contribute to the Formative e-Assessment of First-Year Engineering Students in Mathematics Courses. RUSC, 9:354-370. https://doi.org/10.7238/rusc.v9i1.1277

[16] Clarke J., Dede C. (2010). Assessment, technology, and change. Journal of Research in Teacher Education 42:309-328. https://doi.org/10.1080/15391523.2010.10782553

[17] Erstad O. (2008). Changing assessment practices and the role of ICT. In: J. Knezek and J. Voogt (Eds.), International handbook of information technology in education, Springer, NewYork, NY. pp. 163-180.

[18] Sangwin C.J., Jones I. (2017). Asymmetry in student achievement on multiple-choice and constructed-response items in reversible mathematics processes. Educational Studies in Mathematics, 94:205-222. https://doi.org/10.1007/s10649-016-9725-4

[19] Haapasalo L. (2013). Adapting Assessment to Instrumental Genesis. International Journal for Technology in Mathematics Education, 20:87-93.

[20] Swan M., Burkhardt H. (2012). Designing assessment of performance in mathematics. Educational Designer 2:1–41.

[21] Ferrão M. (2010). E-assessment within the Bologna paradigm: evidence from Portugal. Assessment & Evaluation in Higher Education, 35:819-830. https://doi.org/10.1080/02602930903060990

[22] Sangwin C.J., Köcher N. (2016). Automation of mathematics examinations. Computers and Education, 94:215-227. https://doi.org/10.1016/j.compedu.2015.11.014

[23] Pacheco-Venegas N.D., López G., Andrade-Aréchiga M. (2015). Conceptualization, development, and implementation of a web-based system for automatic evaluation of mathematical expressions. Computers & Education, 88:15-28. https://doi.org/10.1016/j.compedu.2015.03.021

[24] Htwe T.T., Sabaridah I., Low G.K. (2014). Comparative assessment of students' performance and perceptions on objective structured practical models in undergraduate pathology teaching. Singapore Medical Journal, 55:502-505. https://doi.org/10.11622/smedj.2014121

[25] Ristov S., Gusev M., Armenski G. (2015). Massive Development of E-Testing Questions. International Journal of Emerging Technologies in Learning (iJET), 10:46-53. https://doi.org/10.3991/ijet.v10i4.4688

[26] Medland E. (2016). Assessment in higher education: drivers, barriers, and directions for change in the UK. Assessment & Evaluation in Higher Education, 41:81-96. https://doi.org/10.1080/02602938.2014.982072

[27] Reisel J.R., Jablonski M., Hosseini H., Munson E. (2012). Assessment of factors impacting success for incoming college engineering students in a summer bridge program. International Journal of Mathematical Education in Science and Technology, 43:421-433. https://doi.org/10.1080/0020739X.2011.618560

[28] Faridhan Y.E., Loch B., Walker L. (2013). Improving retention in first-year mathematics using learning analytics. 30th Ascilite Conference, Macquarie University, December 1-4 2013, Sydney, Australia. pp. 278-282.

[29] Mišútová M., Mišút M. (2015). Assessment with Information Technology Support. Procedia - Social and Behavioral Sciences, 177:300-305. https://doi.org/10.1016/j.sbspro.2015.02.344

[30] Mišútová M., Mišút M. (2015). ICT as a Mean for Enhancing Flexibility and Quality of Mathematical Subjects Teaching. Lecture Notes in Electrical Engineering, 313:263-267. https://doi.org/10.1007/978-3-319-06773-5_35

[31] Asshaari I., Tawil N.M., Othman H., Ismail N.A., Nopiah Z.M., Zaharim A. (2012). The Importance of Mathematical Pre-University in First Year Engineering Students. Procedia - Social and Behavioral Sciences, 60:372-377. https://doi.org/10.1016/j.sbspro.2012.09.393

[32] Weatherby D.W. (2001). Comparative effects of mathematics intervention strategies on minority engineering students' success. Ph.D. Thesis, Auburn University, Ann Arbor, USA. 96 p.

[33] Drijvers P. (2013). Digital technology in mathematics education: Why it works (or doesn't). PNA, 8:1-20.

[34] Albano G., Ferrari P.L. (2008). Integrating Technology and Research in Mathematics Education: The Case of E-Learning. In: F. J. García-Peñalvo (Ed.), Advances in E-Learning: Experiences and Methodologies, Information Science Reference (an imprint of IGI Global), Hershey, PA, USA. pp. 132-149. https://doi.org/10.4018/978-1-59904-756-0.ch008

[35] Hieb J.L., Ralston P.A.S. (2010). Tablet PCs in engineering mathematics courses at the J.B. Speed School of Engineering. International Journal of Mathematical Education in Science and Technology, 41:487-500. https://doi.org/10.1080/00207390903477467

[36] Lim L.L., Thiel D.V., Searles D.J. (2012). Fine tuning the teaching methods used for second-year university mathematics. International Journal of Mathematical Education in Science and Technology, 43:1-9. https://doi.org/10.1080/0020739X.2011.582171

[37] Tolley P.A., Blat C., McDaniel C., Blackmon D., Royster D. (2012). Enhancing the Mathematics Skills of Students Enrolled in Introductory Engineering Courses: Eliminating the Gap in Incoming Academic Preparation. Journal of STEM Education, 13:74-86.

[38] Mišútová M., Mišút M. (2012). Impact of ICT on the quality of mathematical education. 6th International Multi-Conference on Society, Cybernetics, and Informatics, IMSCI 2012, International Institute of Informatics and Systemics, July 17-20 2012, Orlando, FL, USA. pp. 82-86.

[39] Mišút M., Mišútová M. (2013). Evaluation of ICT Implementation into Engineering Education. International Conference on Advances in Information Technology (ICAIT 2013), IERI, December 27-28 2013, Jeju Island, Korea. pp. 260-265.

[40] Šorm M. (2013). SANCHO Technical Documentation. IS4U, Brno, Czech Republik. 352 p.

[41] Suzuki M., Terada Y., Kanahori T., Yamaguchi K. (2015). New Tools to Convert PDF Math Contents into Accessible e-Books Efficiently. Studies in health technology and informatics, 217:1060-4.

10   Authors

Martin Mišút is a Slovak computer scientist and Associated Professor at Department of Applied Informatics, Faculty of Economic Informatics, University of Economics in Bratislava, Dolnozemska cesta 1/b, 852 35 Bratislava, Slovakia.

Mária Mišútová is a Slovak mathematician and Associated Professor at Institute of Applied Informatics, Automation, and Mathematics, Faculty of Materials Science and Technology in Trnava, Slovak University of Technology in Bratislava, Hajdoczyho 1, 917 24 Trnava, Slovakia.