Beseda, J., Rohlíková, L. (Eds.). (2021).

DisCo 2021: Active Learning in Digital Era: How Digital Tools promote a Conscious, Open-minded, Creative and Social-Oriented Thinking

16th conference reader. Prague: Centre for Higher Education Studies.

DisCo 2021 conference was organized in Prague, Sep 6-7 2021

Publisher: Centre for Higher Education Studies, Jankovcova 933/63, Prague 7, 17000, Czech Republic, www.csvs.cz, www.disconference.eu csvs@csvs.cz, info@disconference.eu

ISBN: 978-80-86302-88-1.

List of reviewers

Prof. Dele Braimoh, ex UNISA, South Africa

Mar Camacho Martí, Ph.D. Universitat Rovira i Virgili in Tarragona, Spain

PhDr. Branislav Frk, Ph.D., University of Prešov, Slovakia

Francesco Pisanu, Ph.D., IPRASE, Italy

Prof. Petra Aczél, Corvinus University of Budapest, Hungary

PhDr. Lucie Rohlíková, Ph.D., Western Bohemia University, Czech Republic

Monika Frania, Ph.D., University of Silesia in Katowice, Poland

Miguel Gea, Ph.D., University of Granada, Spain

Mari Carmen Caldeiro Pedreira, Ph.D., Public University of Navarra, Spain

Dr. Mohammed Ibrahim Alhojailan, Ph.D., DMU University, CCSR. United Kingdom/Saudi University, Saudi Arabia

Mgr. et Mgr. Jan Beseda, Ph.D. Centre for Higher Education Studies, Czech Republic

Mgr. Hana Ovesleová, Ph.D., Western Bohemia University, Czech Republic

Mgr. Jakub Štogr, Navreme Boheme, s.r.o.,Czech Republic

Robert Aust, Universität Leipzig, Germany

Miltiadis D. Lytras, Ph.D., The American College of Greece, Greece

Mgr. Linda Pospíšilová, University of Pardubice

Paraskevi Papadopoulou, Ph.D., The American College of Greece, Greece

Anastasia Misseyanni, Ph.D., The American College of Greece, Greece

Mgr. Veronika Neničková, Masaryk University, Czech Republic

Christina Marouli, Ph.D., The American College of Greece, Greece

María Graciela Badilla Quintana,Ph.D., Universidad Católica de la Santísima Concepcion, Chile

MSc. Lynette Naidoo, UNISA, South Africa

Emilda Roseni, Ph.D., Aleksandër Moisiu University, Durrës, Albania

Mgr. Petr Sudický, Masaryk University, Czech Republic

Mgr. Štěpánka Hronová, MBA, University of Finance and Administration, Czech Republic

Mgr. Tereza Havránková, Western Bohemia University, Czech Republic

Prof. Maurizio Gentile, Ph.D. LUMSA University of Rome/Sapienza University Rome, Italy

Yulia Krylova-Grek, Ph.D. State University of Telecomunnications, Ukraine

Orthena Kotherja, Ph.D., Universiteti „Aleksandër Xhuvani, Albania

Cover graphics Hana Ovesleová

Publication is not a subject of language check. All papers passed a blind review process.

Content

USING E-TEST IN ONLINE MATHEMATICS EDUCATION DURING A PANDEMIC FOR FERENC RAKOCZI II TRANSCARPATHIAN HUNGARIAN COLLEGE OF HIGHER EDUCATION STUDENTS

Gabriella Papp

University of Debrecen, Debrecen, Hungary
p.gabica.17@gmail.com

Abstract

The COVID-19 pandemic has greatly changed mathematics education for the Ferenc Rakoczi II Transcarpathian Hungarian College of Higher Education students. One of the difficult points for a teacher is the survey during online education. In mathematics, open-ended worksheets are typically preferred for summative assessment of students 'knowledge. Although the educational application of e-tests was already present, with the introduction of distance learning it became even more relevant in mathematics. Meanwhile, the e-tests currently measured by teachers are not examined in terms of goodness indicators (objectivity, validity, reliability). In this article, we approach e-tests to measure knowledge of mathematics education at the Ferenc Rakoczi II Transcarpathian Hungarian College of Higher Education. The aim of the research is to present the statistical evaluation of the summary e-test I prepared. I created this on a platform that is easy to assign as there is no need to register students. To preserve objectivity, I created the items according to the dimensions of the Bloom taxonomy by observing the necessary application of open and closed task types. I calculated its reliability using Cronbach's alpha, the low value of which may have been due to the low mean correlation. Because the reliability of the test is low, its validity is also relatively low, as it does not measure true value well.

Keywords

Distance learning. E-test. Mathematics education. Goodness indicator. Bloom-taxonomy.

INTRODUCTION

Due to the COVID-19 pandemic, it was also necessary to introduce online education in the 2020/21 school year. As a teacher, I strive for objective evaluation. In distance learning, I solved the examination and the summative assessment of the level of knowledge with the help of video calls and e-tests. Thus, the problem arose of how to correctly create tests for students knowledge level measurement. In the following, I present the most important information needed for test preparation, which I studied and used for my research and self-examination.

Distance learning and e-test

It is important to understand what is meant by 'distance learning.' Because the technology is evolving, the definition of what distance learning is continues to change (King et al., 2001). According to Keegan (1980), the main elements of a definition of distance education are:

- the separation of teacher and learner which distinguishes it from face-to-face lecturing;
- the influence of an educational organization which distinguishes it from private study;
- the use of technical media, usually print, to unite teacher and learner and carry the educational content of the course;
- the provision of two-way communication so that the student may benefit from or even initiate dialogue;
- the possibility of occasional meetings for both didactic and socialization purposes;
- the participation in an industrialized form of education (Saykılı, 2018).

Distance education is a planned learning experience or method of instruction characterized by quasi-permanent separation of the instructor and learner(s). Such variation includes the types of media or technology used (print, radio, computer); the nature of the learning (workshop, seminar, degree program, supplement to traditional classroom, levels of support); institutional settings; topics addressed; and levels of interactivity support (face-to-face, online, blended, none) (Burns, 2011).

The practices, philosophies and cultures of the persons attempting and developing open and distance learning have impacted how it is designed and conducted (Saykılı, 2018). One of the fastest-evolving modes of distance education is Web-based or online learning (also referred to as virtual learning or e-learning). Indeed, in many parts of the globe, online learning equals distance education. The potential of online learning rests on its ability to do the following:

- Deliver multichannel instruction encompassing print, audio, visual, and video-based content.
- Provide multiple formats for text-based, audio, and video-enabled real-time communication and collaboration with peers across the globe.
- Offer “anytime, anyplace” learning, provided learners have access to the Internet (Burns, 2011).

According Korenova Therefore we can define the term “e-test” dually: 1. In a narrower meaning, the e-test is an electronically controlled didactic test with an option to enrich it with multimedia elements. 2. In a wider meaning, the e-test is an electronic interactive material based on a system of questions and searching for answers created not only for measuring, but also for reaching educational goals (hence can serve as tools for innovative teaching methods). Using e-test we are able not just to determine the students’ knowledge, but with these new digital tools we can increase the students’ motivation, use them during repetition, exercise, controlled discovery methods. The e-test is very attractive from the students’ point of view because the digital world is very close to them (Korenova, 2013).

The use of digital technologies in teaching, distance learning and e-tests is addressed by several authors in publications such as (Guncaga, Korenova & Kostrub, 2018; Korenova, 2015; Korenova & Hvorecky, 2018; Kónya & Kovács, 2019; Slaninka & Simonka, 2017; Vaclavik, Sikorova & Barot, 2019).

Goodness indicators and coefficient alpha

Measurement is central to the construction of a quality student assessment even in the case of a classroom-designed or non-standardized assessments instrument (Eluwa et al., 2011). One of the most important priorities of the automatic e-test assessment and organization is resource saving (Sokolova & Totkov, 2005). The tests used in most tests are usually performance tests, in which case the answer to each item can be correct or incorrect. Such items and variables are called dichotomous (bivalent) variables (Hidegkuti-Balázs, 2015).

In evaluating the quality of an assessment tool, a discussion of reliability and validity is essential. The reliability is the degree to which an instrument consistently measures the ability of an individual or group while validity is the degree to which an instrument measures what it is intended to measure. The classical test theory provides a very simple way of determining the validity and reliability of a test (Eluwa et al., 2011).

In 1951, Cronbach proposed an alpha coefficient (Hidegkuti-Balázs, 2015). Cronbach's alpha is a measure used to assess the reliability, or internal consistency, of a set of scale or test items. In other words, the reliability of any given measurement refers to the extent to which it is a consistent measure of a concept, and Cronbach's alpha is one way of measuring the strength of that consistency (Goforth, 2015).

Figure 1 shows the Cronbach's formula, where n is the number of items, V_t is the variance of the total scores and V_i is the variance of the item scores (Panayides 2013).

Figure 1: Cronbach's alpha.

$$\alpha = \frac{n}{n-1}\left(1 - \frac{\sum_i V_i}{V_t}\right)$$

(Source: Panayides, 2013)

Panayides (2013) describes alpha that has the following important properties: (a) α is the mean of all possible split-half coefficients (b) α is the value expected when two random samples of items from a pool like those in the given test are correlated (c) α is the lower bound of the coefficient of precision … (d) α estimates, and is the lower bound to the proportion of test variance attributable to common factors among the items … (e) α is an upper bound to the concentration in the test of the first factor among the items.

The resulting α coefficient of reliability ranges from 0 to 1 in providing this overall assessment of a measure's reliability. Although the standards for what makes a "good" α coefficient are entirely arbitrary and depend on your theoretical knowledge of the scale in question, many methodologists recommend a minimum α coefficient between 0.65 and 0.8 (or higher in many cases); α coefficients that are less than 0.5 are usually unacceptable, especially for scales purporting to be unidimensional (Goforth, 2015).

If the number of items is small or the average correlation is low, the Cronbach's alpha value will also be low. Nor does a high Cronbach's alpha mean that test items measure a same dimension (Hidegkuti-Balázs, 2015).

Examination using e-Learning environments guarantees reliability, objective assessment, and application of identical assessment criteria for each examinee. Using e-tests teachers can check the knowledge and skills in many more domains compared to the

classical form of examination. Moreover, learners can see outcomes of their achievements and learning progress much faster compared to the time necessary for the classical way of examination (Sokolova & Totkov, 2005).

Bloom-taxonomy

Test questions and assignments, which are included in a concrete e-test can be chosen based on different principles and rules. Opinions of different authors expressed in the literature, are very contradictory. Some authors consider that the test assignments must be chosen according to their type (according to the appropriate classification), and others think that they must be chosen according to their content (including relationship with the subject domain) and/or cognitive objectives of the learning process (Sokolova & Totkov, 2005).

Bloom's Taxonomy is a classification of the different objectives and skills that educators set for their students (learning objectives). The taxonomy was proposed in 1956 by Benjamin Bloom, an educational psychologist at the University of Chicago (Shabatura, 2013). A group of cognitive psychologists, curriculum theorists and instructional researchers, and testing and assessment specialists published in 2001 a revision of Bloom's Taxonomy with the title A Taxonomy for Teaching, Learning, and Assessment (Armstrong, 2010).

Table 1 shows two-dimensional cross-classification of Types of Knowledge by cognitive processing skill.

Table 1: The Taxonomy table.

The Knowledge Dimension	The Cognitive Process Dimension					
	1. Remember	2. Understand	3. Apply	4. Analyze	5. Evaluate	6. Create
A. Factual knowledge						
B. Conceptual knowledge						
C. Procedural knowledge						
D. Metacognitive knowledge						

(Source: Anderson, Krathwohl et al., 2001)

These 6 levels can be used to structure the learning objectives, lessons, and assessments of your course:

- Remembering: Retrieving, recognizing, and recalling relevant knowledge from long-term memory.
- Understanding: Constructing meaning from oral, written, and graphic messages through interpreting, exemplifying, classifying, summarizing, inferring, comparing, and explaining.
- Applying: Carrying out or using a procedure for executing, or implementing.

- Analyzing: Breaking material into constituent parts, determining how the parts relate to one another and to an overall structure or purpose through differentiating, organizing, and attributing.
- Evaluating: Making judgments based on criteria and standards through checking and critiquing.
- Creating: Putting elements together to form a coherent or functional whole; reorganizing elements into a new pattern or structure through generating, planning, or producing (Shabatura, 2013).

In the revised taxonomy, knowledge is at the basis of these six cognitive processes, but its authors created a separate taxonomy of the types of knowledge used in cognition:

- Factual Knowledge
  - Knowledge of terminology
  - Knowledge of specific details and elements
- Conceptual Knowledge
  - Knowledge of classifications and categories
  - Knowledge of principles and generalizations
  - Knowledge of theories, models, and structures
- Procedural Knowledge
  - Knowledge of subject-specific skills and algorithms
  - Knowledge of subject-specific techniques and methods
  - Knowledge of criteria for determining when to use appropriate procedures
- Metacognitive Knowledge
  - Strategic Knowledge
  - Knowledge about cognitive tasks, including appropriate contextual and conditional knowledge
  - Self-knowledge (Armstrong, 2010).

METHODS

During the distance learning introduced due to covid-19, I took several tests in the 2020/21 school year. At first, I didn't consider test theory yet, then I felt it was lacking and I was increasingly trying to create appropriate knowledge level gauges.

I graduated a pilot research at the Ferenc Rakoczi II Transcarpathian Hungarian College of Higher Education Mathematics and Informatics students. The group of 8 people wrote several tests in the spring of 2021, which I had created on several surfaces. In this article, I present the 20-item examination test, the open and closed tasks of which I created based on the Bloom taxonomy. The distribution of taxonomies and items is shown in Table 2.

Table 2: Taxonomies and items.

Bloom's taxonomy	Piece		Items	Piece
B1	5		One choice	6
B2	2		Multiple choice	1
B3	2		True – false	4
C1	4		Matching	5
C2	2		Input digit	2
C3	5		Input text	2
Sum	**20**		**Sum**	**20**

(Source: Own)

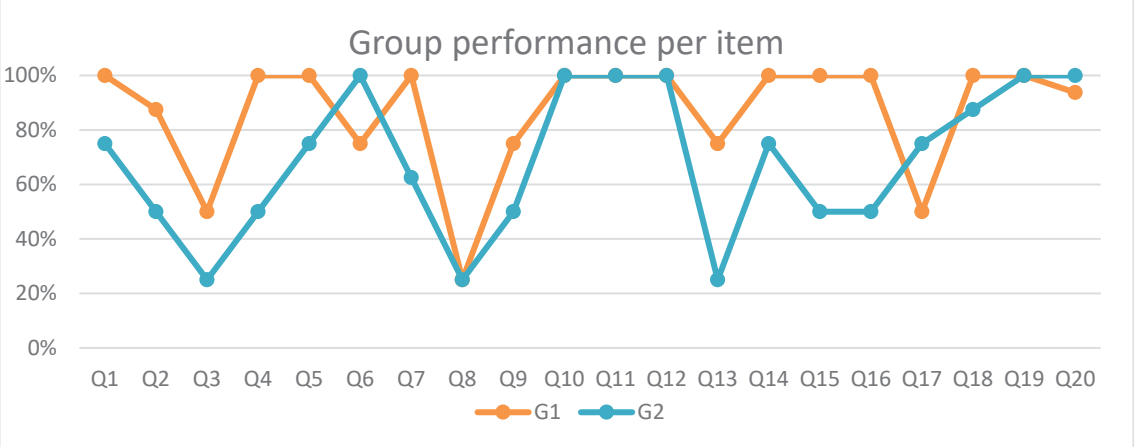
I used the Online Test Pad (https://onlinetestpad.com) platform to create the test. I chose this interface for several reasons: here I found most of the item types I was looking for with a simple editing interface; complex configuration options in terms of filler constraints; the test can be distributed by sharing a link without registering and creating a group (group creation is also possible on the interface if necessary); the test downloadable in pdf format for offline application; can be easily completed online using both a computer and a mobile phone after opening the link.

Students had 30 minutes to complete the test, while they were video call with me. On the video call I saw that they used only their knowledge. In addition to theoretical knowledge, it was also necessary to solve practical tasks.

RESULTS

In the summative e-test, all students managed to reach the required minimum of 50% completion. The best result was 95%. In order to analyse the effectiveness of the e-test more accurately, I classified the 8 results received into 2 classes in the group above (G1) and below (G2) the performed average. Figure 2 shows well which items (Q1, Q2,…, Q20) show surprising results.

Figure 2: Results of groups per item.



(Source: Own)

Let us first consider tasks Q6 and Q17, for which several correct answers were received from group G2. Question Q6 based on Bloom taxonomy C3 (procedural + apply) was an open-ended question, students needed defining the limit of a function. I ruled out the possibility of choice, it was necessary to perform an independent calculation, so better

results could be expected from above-average students. For problem Q17, the solution of a B1 (conceptual + remember) true-false type item had to be given. Here, even in the case of incorrect and correct solutions, in addition to remember, there are opportunities for uncertainty and non-remember in both groups, which led to the results seen.

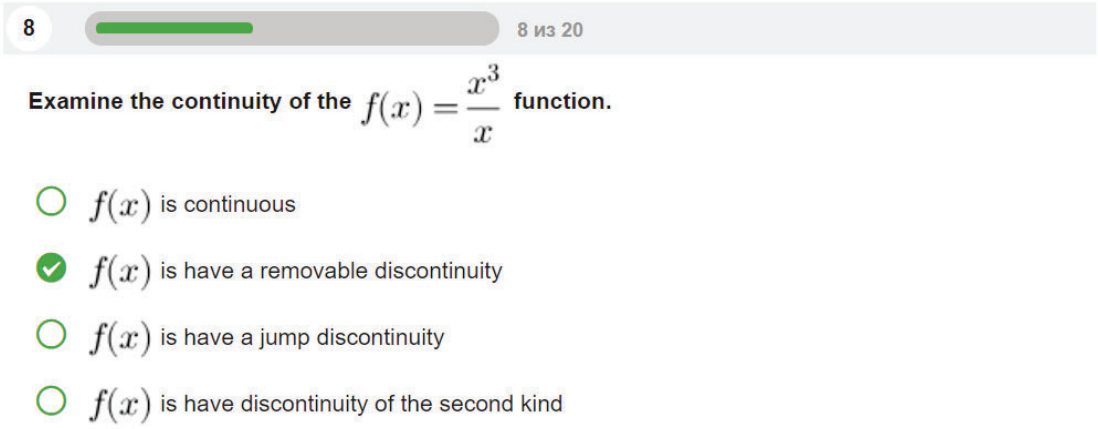
Questions Q10, Q11, Q12, and Q19 were answered correctly by all members of both groups. These questions were on Bloom B1 and C1 (procedural + remember) level and they were matching, input text, true-false, and one choice type of items. It can also be considered a mistake of the research that these items had to be recalled for basic knowledge, although I paid close attention to the students to give a correct answer only with their clear knowledge.

Statistically, the most difficult questions were questions Q3, Q8, and Q13, in which the groups gave correct or incorrect answers while the results remained at the level appropriate to the group or performed equally poorly. All three items contained practical assignments, based on Bloom taxonomy C3 (procedural + apply) and one choice type.

In question Q3, the domain of a rational function had to be specified correctly, for which the criterion of the square root in the denominator of the fraction had to be known and applied. Under the square root was the quadratic equation could be easily solved using the Vieta's formulas.

Question Q8, shown in Figure 3, proved to be the most difficult item for both groups. One correct answer was received to this question in which the continuity or discontinuity of a function had to be determined.

Figure 3: Item 8 for the test.



(Source: Own)

In this case, I assume that due to lack of time, students did not count but chose one answer from those listed. I drew this conclusion because in the Q7 question before it, the formulas had to be matched to the same concepts in Q8. Everyone in group G1, while only one student in G2 paired correctly, the others could only solve half of the question. In group G2, not the student who gave correct answer to Q7 gave the correct answer to Q8.

For question Q13, the difference between the 2 groups was greater. While 1 person from group G1 spoiled the answer, 1 person from group G2 was able to answer correctly. In the task, the primitive of the given function had to be marked. To get the right answer, the students had to know the concept of a primitive function and the primitive of

trigonometric functions. A student who confuses the concepts of primitive and derivative of a function had the opportunity to mark a wrong answer.

Surprisingly many students made a mistake in these questions in both groups. This allows me a condition to assume that in these cases, the multiple-choice type allowed room for inaccurate work or wrong choices without calculations and thinking.

The questions not mentioned earlier show the expected difference between the two groups, which suggests that they are good in reliability and well measured. These include a mix of item types and questions corresponding to Bloom taxonomy levels. Nevertheless, from the point of view of statistical analysis, the examined task set cannot be considered a reliable e-test, because the Cronbach's alpha value of 0.15 is too low, so it falls into the unacceptable category.

DISCUSSION AND CONCLUSION

Seeing my results at the end of the analysis, I conclude that while the best start was to delve in the theoretical knowledge before trying it out in practice, I don't have to stop this if the expected results don't show up on the first try.

One of the aims of the research I achieved by making it easier to create an objective e-test using the Bloom taxonomy. It was helped to maintain objectivity by being created on an online platform and improved it by the system itself. Its validity depends on its reliability, which, given the results obtained, I cannot yet establish with certainty.

Figure 2 suggests that the e-test examined, and its results are rather reliable. In contrast to the low value of Cronbach's alpha. This may have been due to low mean correlation or small sample size. To clarify this, it would be advisable to repeat the research, preferably in a larger group, as more accurate calculations can be performed with a multi-item sample. If I get a similar result after repeating, I can conclude that this e-test is unreliable.

Overall, the research has not been fully effective, so further research is needed.

REFERENCES

Anderson, L. W., Krathwohl. D. R. (2001). *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives.* Available at https://www.uky.edu/~rsand1/china2018/texts/Anderson-Krathwohl%20-%20A%20taxonomy%20for%20learning%20teaching%20and%20assessing.pdf (Accessed 24. June 2021).

Armstrong, P. (2010). Bloom's Taxonomy. *Vanderbilt University Center for Teaching*. Available at: https://cft.vanderbilt.edu/guides-sub-pages/blooms-taxonomy/ (Accessed 24. June 2021)

Burns, M. (2011). Distance Education for Teacher Training: Modes, Models, and Methods. *Education Development Center, Inc. Washington, DC*. Available at: https://www.edc.org/sites/default/files/uploads/Distance-Education-Teacher-Training.pdf (Accessed 20. June 2021).

Eluwa, O. I., Eluwa, A. N., Abang, B. K. (2011). Evaluation of Mathematics Achievement Test: A Comparison Between Classical Test Theory (CTT) and Item Response Theory (IRT). *Journal of Educational and Social Research*, 1(4), 99-106.

Goforth, C. (2015). Using and Interpreting Cronbach's Alpha. *University of Virginia Library*. Available at: https://data.library.virginia.edu/using-and-interpreting-cronbachs-alpha/ (Accessed 21. June 2021).

Guncaga, J., Korenova, L., & Kostrub, D., (2018.) The educational research focused on the development of mobile technologies in education. In M. Artois, *Teaching with technology: Perspectives, challenges and future directions*, NOVA Science Publisher, New York, 57-115.

Hidegkuti, I., Balázs, K. (2015). Tesztelmélet. In: Balázs, Kovács, J., & Münnich, Á. *Pszichológiai Módszertani Tanulmányok*. Debreceni Egyetemi Kiadó, 65-95.

King, F. B., Young, M. F., Drivere-Richmond, K., & Schrader, P. G. (2001). Defining distance learning and distance education. *AACE Review*, 9(1).

Korenova, L. (2013). Usage possibilities of e-tests in a digital mathematical environment. *Usta ad Albim BOHEMICA*, 13(3), 78-83.

Korenova, L. (2015). What to use for mathematics in high school: PC, tablet or graphing calculator? *International Journal for Technology in Mathematics Education*, 22(2), 59-64.

Korenova, L., & Hvorecky, J. (2018.) Applying QR Codes in: Facilitating Mathematics and Informatics Education. *International Journal of research in E-learning.* 4(2), 33-44.

Kónya, E., & Kovács, Z. (2019). Do Calculators Support Inductive Thinking? *The Electronic Journal of Mathematics and Technology*, 13(2), 181-189.

Kovács, Z., Recio, T., & Vélez, M. P. (2018). Using automated reasoning tools in GeoGebra in the teaching and learning of proving in geometry. *Journal for Technology in Mathematics Education*, 25(2), 33-51.

Panayides, P. (2013). Coefficient Alpha: Interpret With Caution. *Europe's Journal of Psychology*, 9(4), 687–696.

Saykılı, A. (2018). Distance education: Definitions, generations, key concepts and future directions. *International Journal of Contemporary Educational Research*, 5(1), 2-17.

Slaninka, F., & Simonka, Z. (2017). 'Probability distributions in maxima software', in *4. ročník konference MITAV, Fakulta vojenských technologií Univerzity obrany v Brně*.

Shabatura, J. (2013). *Using Bloom's Taxonomy to Write Effective Learning Objectives*. Available at: https://tips.uark.edu/using-blooms-taxonomy/ (Accessed 24. June 2021).

Sokolova M., Totkov, G. (2005). 'About Test Classification in E-Learning Environment', in: *International Conference on Computer Systems and Technologies*, pp. 211-216.

Vaclavik, M., Sikorova, Z., Barot, T. (2019). 'Approach of Process Modeling Applied in Particular Pedagogical Research', in *Cybernetics and Automation Control Theory Methods in Intelligent Algorithms, Advances in Intelligent Systems and Computing*. Springer, pp. 97–106.

Václavík, M. et al. (2018). 'Particular Analysis of Normality of Data in Applied Quantitative Research', in *2nd Computational Methods in Systems and Software 2018: Computational and Statistical Methods in Intelligent Systems, Advances in Intelligent Systems and Computing*, pp. 353-365.