EMBEDDED ASSESSMENT THROUGHOUT THE PROBLEM-BASED LEARNING – A PATH TO AUTHENTIC LEARNING EXPERIENCES[*]

Borislava Kirilova

Faculty of Mathematics and Informatics,
Sofia University "St. Kliment Ohridksi", Sofia, Bulgaria
e-mail: b.kirilova@fmi.uni-sofia.bg

This article examines the benefits and challenges associated with Problem-Based Learning (PBL) as a teaching and assessment tool, a constructivist approach that promotes a more pragmatic and student-centered mode of assessment. A tangible illustration of the proposed assessment tool is presented. The study aims to analyze different aspects of mathematical competence in line with the European Commission Framework for Lifelong Learning, such as critical thinking, practical application in authentic problems, and more effective knowledge acquisition. Data were collected through direct classroom observations, open-ended questionnaires, surveys, and personal interviews. The methodology was applied to a sample of 60 students from American College of Sofia (ACS) who participated in project work instead of conventional testing. The results of the study revealed that assessing students' progress through problem-based tasks not only enhances the development of students' mathematical competencies compared to traditional methods, but also fosters lifelong learning skills. However, this approach requires a more significant time commitment compared to traditional testing.
Keywords: Problem-Based Learning, embedded assessment, authentic assessment, binomial distribution, normal distribution, hypotheses.

ВГРАДЕНО ОЦЕНЯВАНЕ ПО ВРЕМЕ НА ПРОБЛЕМНО-БАЗИРАНО ОБУЧЕНИЕ – ПЪТ КЪМ АВТЕНТИЧЕН УЧЕБЕН ОПИТ

Борислава Кирилова

Факултет по математика и информатика,
Софийски университет „Св. Климент Охридски", София, България
e-mail: b.kirilova@fmi.uni-sofia.bg

В статията се разглеждат предимствата и предизвикателствата, свързани с проблемно-базираното обучение (ПБО) като инструмент за преподаване и оценяване – конструктивистки подход, който насърчава по-прагматичен и ориентиран към учениците способ за оценяване. Представена е конкретна илюстрация на предложения инструмент за оценяване. Изследването има за цел да анализира различни аспекти на математическата компетентност в съответствие с Европейската рамка за учене през целия живот, като критично мислене, практическо приложение на придобитите знания при решаване на проблеми от реалния живот и умения за учене. Данните са събрани чрез преки наблюдения в класната стая, въпросници с отворен отговор, анкети и лични интервюта. Методологията беше приложена върху извадка от 60 ученици от Американски колеж в София, които участваха в работа по проект вместо в конвенционален тип изпитване. Резултатите от изследването разкриват, че оценяването на напредъка на учениците чрез проблемно базирани задачи не само повишава развиването на математическите им компетентности, в сравнение с традиционните методи, но и възпитава умения за учене през целия живот. Този подход обаче изисква по-значителен ангажимент откъм време спрямо традиционното изпитване.

Ключови думи: проблемно-базирано обучение, вградено оценяване, автентично оценяване, биномно разпределение, нормално разпределение, статистически хипотези.

Introduction. In the realm of education, the conventional test has consistently served as the main tool for assessing the knowledge and skills acquired during the educational process. However, as educators have become more steeped in the constructivist model of learning, there is a growing consensus that traditional examinations are not the only instrument of assessing knowledge. Constructivism requires to abandon traditional assessment, standardized tests, and checklists. In this paradigm, assessment is seamlessly integrated into the learning process, giving students a more meaningful role in assessing their own progress [19]. Traditional test papers, characterized by limited time for work and covering specific learning material, in most cases measure mainly the memorization and the repetition of facts [8]. There is a significant likelihood that certain cases do not precisely portray the student's capacity to apply knowledge in practical situations or think critically about authentic problems. Furthermore, the uniform nature of examinations, designed to fit all, fails to consider diverse learning styles and may result in a distorted representation of a student's true capabilities. Papers can cause stress and anxiety in students, which can interfere with their performance and not be indicative of their actual mathematical competencies [20]. Problem-Based Learning (PBL) is an instructional approach that integrates various forms of active learning opportunities, offering students the chance to collaborate and engage in a low-stress environment. In a study conducted by Chapell et al. [4], it was found that students experiencing high anxiety exhibited significantly weaker performance on a time-pressure test compared to their peers with lower anxiety levels. This outcome was attributed to the impact of stress and anxious thoughts, which were observed to reduce working memory capacity [12]. Project-based assessment provides authentic learning experiences that extend beyond the classroom and mere academic content. However, it is important to consider the fact that standardising the assessment of projects and eliminating subjectivity present challenges that significantly reduce their advantages over examinations. Assessment conducted under the guidance of a proctor provides a fairer grading process for students

as it reduces the opportunity to present someone else's work as their own [17]. Respectively, a complete transition to the implementation of assessment in the current learning process could eliminate the advantages of problem-based learning. On the other hand, if our attention is solely directed toward assessments conducted in the classroom under the direct supervision of the teacher, we should anticipate no disparity in the prevalence of cheating between traditional paper-and-pencil tests and computer-based grading [13].

Methodology. The initial stage of implementing project-based evaluation requires careful selection of the topics to be addressed in line with the learning objectives. To foster the development of mathematical competencies [20], it is essential to select topics that are relevant to real-world challenges, interdisciplinary in nature, and require research, data analysis, creativity, problem solving, and critical thinking. Creating transparent and objective assessment criteria is crucial for fostering student motivation [5]. Rubrics should be formulated to evaluate not only the final result but also the process of engagement, collaboration, and individual contributions. Conventional assessment methods often do not meet these requirements. In PBL, students actively participate in classroom learning and demonstrate their progress while practicing the appropriate type of tasks or developing problem-solving skills. This method provides numerous opportunities for authentic, embedded assessment. Embedded assessment means that the assessment of students' progress and achievement is seamlessly integrated into ongoing practical activities, whereas non-embedded assessment takes place outside the regular learning process, often in the form of tests. Assessing student progress and performance during a PBL session can be effectively achieved through embedded assessment [3]. The embedded assessment tool documents what tutors observed while students were engaged in learning or completing a task. Authentic assessment is associated with criteria defining to what extent the problem addressed is part of real life. Students are required to engage in authentic problems, rather than repeating facts.

Embedded assessment. Analyzing real-world cases by problem-based assessment provides students with exposure to different problem-solving methodologies and enhances their ability to apply these approaches. This helps bridge the gap between theoretical understanding and practical application. The examples presented in this section were developed by ACS students during mathematics classes, with the grades earned replacing a current assessment.

The examples presented in this section were developed by students at the American College of Sofia during mathematics classes, and the resulting assessment replaces a current test grade.

Problem-based assessment for 2021/2022 year: The aim of the project is to create two problems related to the topic "Statistical inference", as follows: 1. Statistical inference with a binomial distribution model on data from a learning test. 2. Statistical inference with a normal distribution model on data from a learning test. Choose a real-life topic of interest to you. It can be related to a personal story; note it in the introduction to your paper. If not, provide a rationale for why you chose this particular topic and how it might apply true-to-life. Before you start actual work on the project, you need to choose a general data set. Each of these sets should contain a sample of at least 100 items from reliable sources. Cite them in the last part of the report. Make sure that none of your classmates have already chosen the same data samples by entering them into a table that is common to the grade level. It is not even acceptable for students from different

classes to use the same data. At the conclusion of the final report, please complete the following teacher-created survey to test how effective you think project-based testing is: Perceived Importance: On a scale of 1 to 5, evaluate the importance of assessments based on authentic problems for your subject matter expertise knowledge (1 = Not important at all, 5 = Extremely important). Knowledge Acquisition: Do you believe that information learned through projects is better retained compared to traditional testing methods (Yes/No)? If yes, please explain why. Application to Real Life: How well do you think the knowledge gained through project work can be applied to real-life situations (Very Poor, Poor, Neutral, Good, Very Good)? Critical Thinking Skills: In order to understand to what extent do projects enhance students' problem-solving skills, the following questions were asked: Do you think that the problem-based work helped to improve your ability to research, extract and evaluate the data collected (Not at all, Slightly, Moderately, Very much, Extremely)? How has project-based testing contributed to your ability to develop solutions with the use of statistical analysis (Little, Somewhat, A lot)? Suggestions for Improvement: What suggestions do you have to improve the effectiveness of project-based testing? Is there anything else you would like to share about your experience with project-based testing?

In the original document, this section presents solved problems similar to those expected to be created by students. However, the author prefers to show students' solutions after formative assessment and remediation were applied. Both of the following topics have their roots in the reality surrounding the students and more particularly the COVID-19 pandemic environment which was so live during the time of the research.

An example of statistical inference with a binomial distribution model proposed by a student at an American College of Sofia. The first one of the problems is related to the effectiveness of the vaccines.

The introduction was as follows: the first approved vaccine passed three stages of clinical trials as required by the US's Center for Disease Control (CDC) and the Federal Drugs Administration (FDA) before finally being approved for emergency use by the FDA on December 11, 2020. It remained the solely approved vaccine until April 23, 2021 when the "Johnson & Johnson" vaccine was approved as well. For the purposes of his work the student has assumed the number of people vaccinated with the "J&J" vaccine is negligible and therefore the effects on the population can also be neglected. In particular to the basic used vaccine – that one developed by "Pfizer", after the third and final stage of their clinical trials the producer reported a 95% (at least) protection rate against vaccine breakthrough infection. This results in a 5% (at most) chance of getting infected with COVID-19 after completing the two doses series of the "Pfizer" vaccine. Meanwhile, the Center of Disease Control (CDC) [21,22] provided data about the total number of vaccinated people who fully completed the Pfizer track and the number of people with the "Pfizer" vaccine who got the infection up to one month after completing the series. The numbers were as follows by April 2021: $n = 10\ 135\ 725$ people got the two jabs from the "Pfizer" and $x = 12\ 611$ got infected afterwards.

Based on the facts above and assuming that the possibility of each vaccinated person to get infected constitutes independent events the student constructed the experiment based on a binomial distribution such as $B(n; \pi) = B(10\ 135\ 725; 0.05)$ and the following hypothesis definitions: the null hypothesis $H_0 : \pi \leq 0.05$ suggesting that there isn't a greater than reported risk of getting the virus after the vaccination. In addition, the

alternative hypothesis is $H_1 : \pi > 0.05$ suggesting that "Pfizer" incorrectly measured the rate of getting the virus after jabbing and the risk is higher than reported by them. Significance level $\alpha = 0.05$ meaning that results will hold with 95% confidence is set by the project guidelines. To solve the issue the student has cleverly applied a straight calculation saying $P(x \geq 12\,611) = 1 - BINOM.DIST(12\,611; 10\,135\,725; 0.05; TRUE) \approx 1 > 0.05 = \alpha$.

Shortly this can be read as that the probability of getting more than the reported infection cases within the observed population and given the reported vaccine efficiency strongly exceeds the assumed confidence level and therefore, we can be 95% confident that we cannot reject the null hypothesis. Furthermore, with 95% confidence, it could be concluded that "Pfizer" did not misreport the rate for possible infections after vaccination and that it is less or equal to 5% with 95% confidence.

An example of statistical inference with a normal distribution model proposed by an American College student. The second problem explores a possible relation between the vaccination levels and the death cases.

The rationale is based in the statistical data for Bulgaria and Denmark and in particular the deaths occurring in the two countries during April 2021. The student choses to compare Bulgaria's death rates to those of Denmark due the last country having the closest, among all other European countries, number of citizens residing in it compared to our home country (Bulgaria – 6 896 663, Denmark $\sim$ 5.84 million). In both countries vaccination begins on December 27, 2020. According to "Our World in Data", however, by April 2021, only 3% of the Bulgarian population was fully vaccinated, whereas in Denmark about 10% were fully vaccinated (and 13% had at least one dose put on) [23, 24].

Based on the information above it is naturally expected that the mean of Bulgarian deaths per day during April 2021 would be greater than that of Denmark. Statistical data for Bulgaria in this particular month summarizes to $(\mu, \sigma) = (110, 47.63)$ and for Denmark $(\mu, \sigma) = (2.2, 1.21)$ for $n = 30$ daily observations. Thus, the following hypothesis definitions are constructed: the null hypothesis (H_0) follows fact that the average deaths per day in the Scandinavian country is $\mu_0 = 2.2$ ($H_0 : \mu \leq \mu_0 = 2.2$) i.e average death cases in Bulgaria will not exceed the average in Denmark, and the alternative hypothesis is that Bulgaria's deaths per day are bigger than those of Denmark ($H_1 : \mu > 2.2$). Project guidelines request a significance level of $\alpha = 0.05$ meaning that results will hold with 95% confidence and therefore the corresponding *z-score used is 1.645*.

Given the above setup the empirical $z_e = 16.64$ is calculated. As $16.64 \gg 1.645$ or $z_e \gg z$ a conclusion was made that the null hypothesis H_0 can be rejected with more than 95% confidence. In addition, with confidence higher than 95% it can be concluded that average death cases for Bulgaria is higher than the average death cases for Denmark.

Results and discussion. In the field of education, the traditional system of assessing students through examinations has long been a subject of debate [11]. Although tests are designed to assess students' knowledge and skills acquired during learning activities, they often fail to capture essential skills such as critical thinking, problem solving, and creativity. Here, an attempt is made to report not only the results based on the final product, but also the work process itself. Perceived Importance: The author is aware that the answers obtained depend more on how well the students have understood that the purpose of this type of testing is to test their real-world problem-solving skills and subject matter competence. Accordingly, answers can be highly subjective. To avoid

inaccuracies, detailed explanations of what is meant by a level choice on the scale were presented on a whiteboard and additional personal interviews were conducted. Following these steps, the final consensus revealed that approximately 92% (rounding to a whole number was used for the purpose of the discussion) of the students believed that exams emphasizing practical skills created the necessary environment to move into a higher level of skill development according to Bloom's taxonomy as revised and specified for educational purposes in mathematics education [1]. Knowledge Acquisition: 87% of students felt that knowledge gained through project work was better remembered in comparison to their preparation for traditional testing methods. The remaining 13% felt that PBL did not suit their learning style because they had to work in an environment that required a high degree of self-direction. This thesis is supported by Prince & Felder [14] who claimed that students who struggle with the level of self-direction and responsibility required in a PBL environment are no exception. Research skills: By creation of hypothesis testing problems 97% of the students felt that skills such as research, data collection, extract, and interpretation were developed. Analytical skills: Working with an actual data encourages students to critically evaluate information, make interdisciplinary connections, and arrive at meaningful conclusions in the context of the problem rather than simply following problems from the textbook. Data analysis often involves using statistical software and data processing tools. Students gain practical experience with technology, enhancing their digital literacy skills, which are crucial in today's information-driven world. According to 92% of students, the competencies developed through the "Statistical inference" project would have benefited them beyond the classroom. Application to Real Life: The use of realistic data often requires students to integrate knowledge from different disciplines – a practice disapproved of by only 23% of the experimental group. These challenges promote lifelong learning as students can apply these interactive models in a variety of contexts throughout their future lives. This interdisciplinary approach fosters a holistic understanding of complex issues, promoting a broader perspective and encouraging students to draw connections between different areas of study. Suggestions for Improvement: Although embedded PBL evaluation offers numerous advantages, it also has some potential disadvantages compared to conventional, non-embedded evaluation. Here are some of the disadvantages students have associated with embedded assessment: Developing and completing a project takes significantly longer than preparing for and taking a traditional exam [10] which was confirmed by 96% of the experimental group. During the interviews, some of the students (12%) shared their concerns about whether PBL adequately prepares them for the State Profile Math Exams. These students prefer traditional exams, since in their opinion, they are more efficient in covering a broad range of content in a short amount of time. Traditional exams often mirror the format of standardized tests more closely [7]. PBL may focus on a specific topic, potentially leaving some other less explored due to the luck of time. Kirschner et al. [9], shared that PBL might not cover a broad range of content efficiently, potentially leaving gaps in students' knowledge compared to traditional exams. Assessing problem-based learning can be subjective and challenging. Evaluation criteria can vary, and it can be difficult to quantify and standardize the evaluation of different projects [2]. According to Thomas [16], project-based learning evaluation can be subjective and challenging. Project assessment can involve more ambiguity and variability in grading compared to the objective nature of traditional ongoing examinations. Furthermore, project management, monitoring and

evaluation involves the periodic collection, processing, and storage of significant amounts of information by the tutors [6]. In contrast, traditional tests may be more familiar and easier to manage for educators without extensive professional practice [15]. Ultimately, the choice between embedded and traditional assessment depends on the educational goals, the nature of the subject, and the preferences of both educators and students. Many modern educational approaches aim to strike a balance between these two methods to provide a well-rounded learning experience. It's important to note that while these criticisms exist, proponents of project assessment argue that many of these challenges can be addressed through thoughtful design, proper training, and ongoing assessment refinement. The effectiveness of PBL often depends on how well it is implemented and integrated into the broader educational context.

CONCLUSION. The requirements of the educational process are evolving and so are the methods of assessment. Although tests continue to be the main tool, the integration of project-based assessment offers an effective alternative. A constructivist approach gives priority to authentic learning experiences, promotes the practical application of knowledge and cultivates key competences for lifelong learning. Although there are challenges in its implementation, the benefits are undeniable – a more committed, adaptive and prepared generation of students for the changing demands of the 21st century.

REFERENCES

[1] L. ANDERSON, D. KRATHWOHL. A Taxonomy for Learning, Teaching, and Assessing. A Revision of Bloom's Taxonomy of Educational Objectives. Longman, 2001.

[2] J. BIGGS, C. TANG, G. KENNEDY. Teaching for quality learning at university. SRHE and Open University Press, 2022. ISBN 13:978-0-33-524275-7.

[3] PH. BLUMBERG. Assessing Students During the Problem-Based Learning (PBL) process. *Medical Science Educator*, **15**, no.2 (2006).

[4] M. CHAPELL et al. Test anxiety and academic performance in undergraduate and graduate students. *Journal of Educational Psychology*, **97**, no. 2 (2005), 268–274, `https://doi.org/10.1037/0022-0663.97.2.268`.

[5] M. FLORES et al. Perceptions of effectiveness, fairness and feedback of assessment methods: A study in higher education. *Studies in Higher Education*, **40**, no. 9 (2014), 1523–1534, `https://doi.org/10.1080/03075079.2014.881348`.

[6] E. HADJIKOLEV et al. Conceptual model of a virtual project management center. Scientific Research of the Union of Scientists in Bulgaria – Plovdiv, series B. Natural Sciences and Humanities, Vol. XVI, 2013, ISSN 1311-9192

[7] J. HATTIE. Visible learning: A synthesis of over 800 meta-analyses related to achievement, 2009. London: Routledge.

[8] T. HATTINGH et al. Student learning behaviours around assessments. *Australasian Journal of Engineering Education*, **24**, no. 1 (2019), 14–24, `https://doi.org/10.1080/22054952.2019.1570641`.

[9] P. KIRSCHNER et al. Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, **41**, no. 2 (2006), 75–86, `https://doi.org/10.1207/s15326985ep4102_1`.

[10] J. MERGENDOLLER et al. The effectiveness of problem-based instruction: A comparative study of instructional methods and student characteristics. *Interdisciplinary Journal of Problem-Based Learning*, **1**, no. 2 (2006), `https://doi.org/10.7771/1541-5015.1026`.

[11] P. Nestorova et al., The test – (un)reliable way of checking knowledge. *Scientific works of UFT*, **59**, (2012) "Food, science, engineering and technologies", 962–965, ISSN 1314-7102.

[12] M. Owens et al. Anxiety and depression in academic performance: An exploration of the mediating factors of worry and working memory. *School Psychology International*, **33**, no. 4 (2012), 433–449, `https://doi.org/10.1177/0143034311427433`.

[13] Ph. Petrov. Prevention of Cheating in Paper-And-Pencil Tests for the Organizational and Technological Conditions of the Bulgarian Schools with the Help of Web-Based Test Generator Software. *Journal "Mathematics, Computer Science and Education"*, **1**, no. 1, (2018).

[14] M. Prince, R. Felder. Inductive teaching and learning methods: Definitions, comparisons, and Research Bases. *Journal of Engineering Education*, **95**, no. 2 (2006), 123–138, `https://doi.org/10.1002/j.2168-9830.2006.tb00884.x`.

[15] K. Simons, P. Ertmer. Scaffolding disciplined inquiry in problem-based learning environments. *The International Journal of Learning: Annual Review*, **12**, no. 6 (2006), 297–306, `https://doi.org/10.18848/1447-9494/cgp/v12i06/47900`.

[16] J. Thomas. A review of research on project-based learning. 2000, California, Autodesk Foundation.

[17] J. Wilkinson. Staff and student perceptions of plagiarism and cheating. *Int. J. Teach. Learn. High. Educ.* **20** (2009), 98–105.

[18] D. William. What Is Assessment for Learning? *Studies in Educational Evaluation*, **37**, no. 1, (2011), 3–14.

NOTES

[19] `http://www.keytttcd.cct.bg/images/files/D8-KeyTTT%20Methodology%20basisBG.pdf`.

[20] `https://www.britishcouncil.org/sites/default/files/youth-in-action-keycomp-en.pdf`.

[21] Data on the Daily Number of New Reported COVID-19 Cases and Deaths by EU/EEA Country." *European Centre for Disease Prevention and Control*, 7 Apr. 2022, `https://www.ecdc.europa.eu/en/publications-data/data-daily-new-cases-covid-19-eueea-country`.

[22] Rates of Covid-19 Cases or Deaths by Age Group and Vaccination Status. *Centers for Disease Control and Prevention*, Centers for Disease Control and Prevention, `https://data.cdc.gov/Public-Health-Surveillance/Rates-of-COVID-19-Cases-or-Deaths-by-Age-Group-and/3rge-nu2a`.

[23] H. Ritchie et al. Coronavirus (COVID-19) Vaccinations. *Our World in Data*, 5 Mar. 2020, `https://ourworldindata.org/covid-vaccinations`.

[24] Vaccine Timeline. Historic Dates and Events Related to Vaccines and Immunization, `https://www.immunize.org/timeline/`.