

Methodological Model of Organizing Practical Training on the Basis of Self-Check Systems

Akhtamova Laziza Azam kizi
Bukhara State Technical University
ORCID: 0009-0004-0219-3552

Abstract - This article describes the development of a methodological model for organizing practical exercises in a digital learning environment based on self-checking systems and its empirical testing. The study was organized in a quasi-experimental design based on the integration of the LMS Moodle and the ACC1 self-checking plugin. The results showed that in the experimental group, the response time to feedback was significantly reduced, the time to success was reduced, and the percentage of task completion increased. The results of the correlation analysis identified the response time to feedback and success as the strongest predictors of learning outcomes, while the number of attempts was a relatively weak factor. It was argued that self-checking systems, when designed not only as an automated assessment tool in practical exercises, but also as a methodological mechanism that controls the formative cycle, can simultaneously increase educational effectiveness and process efficiency.

Keywords: self-checking system, methodological model, automated assessment, formative feedback, resubmission, learning analytics, Moodle, ACC1.

INTRODUCTION

In a digital learning environment, practical exercises are a cornerstone in the formation of student competence. It is in the process of completing a practical assignment that knowledge becomes a “working skill”, errors are diagnosed, and correction strategies are formed. However, the organization of practical exercises in higher education often falls between two conflicting needs. On the one hand, the student needs quick, clear, and understandable feedback on each attempt, and on the other hand, in the conditions of increasing teacher workload and the number of groups, it becomes difficult to provide such feedback individually and regularly for each student. As a result, the stability of assessment criteria decreases and the number of delays in the process increases, the student repeats errors without deep analysis, and the practical exercise becomes a “one-time assignment” rather than “result-oriented iterative learning”. This problem is especially evident in disciplines such as programming, algorithmics, databases, and modeling, since the quality of learning in these disciplines largely depends on how the “error-correction-retest” cycle is organized.

In recent years, self-checking systems have enhanced the capacity to fill this gap in practical training. The system automatically checks the student’s assignment, returns the result, encourages iterative improvement through resubmission, and provides monitoring for the teacher. However, an analysis of international scientific sources shows that many solutions mainly focus on improving technical verification mechanisms, but the issue of their full integration into teaching methods has not been sufficiently formulated in the form of a methodological model. For example, reviews of automated assessment systems for programming assignments highlight features such as resubmission policies, security, and test creation as central differentiators of the systems, but also indirectly indicate that pedagogical management is often left out of the system [1]. Subsequent systematic reviews reinforce this conclusion, namely that automated feedback is often limited to the level of “pass/fail” and “expected-actual result difference”, and does not sufficiently cover didactic requirements such as explaining the cause of the error, directing the student toward correction, and linking it to learning objectives [2]. One of the most recent large systematic reviews also notes that most assessment tools focus on automating correctness, while relatively underestimating quality indicators such as code readability, documentation, and especially methodologically enriched feedback [3]. This means that the issue is not just about “is there/isn’t there automatic verification?”, but also about “how does automatic verification redesign the methodology of practical training?”

The impact of feedback on learning has been widely documented in educational psychology and didactics. For example, formative feedback is a powerful learning factor when it is timely, specific, linked to learning objectives, and can change the student’s subsequent behavior [4]. It has been shown that the content and delivery of feedback, such as pointing out errors, explaining the reasons, guiding students to correct them, and encouraging self-examination, are crucial [5]. Thus, the effectiveness of self-examination systems depends not only on the computational resources or set of tests, but also on how feedback is structured as a methodological scenario and how it is linked to rubrics and competency indicators. International empirical research supports this idea. For example, during the pandemic, when adapting a practical programming course to a fully online format, an automated, student-centered assessment tool was used, which showed positive student experiences and satisfaction, and when such tools were methodologically properly integrated, they were shown to work consistently in conditions of high student-teacher ratios [6]. Similarly, tools tested at the production level show that they are effective in practice, but also imply the need for methodological regulation to stabilize the didactic value of these tools [7].

Another important dimension of assessment in a digital environment is academic integrity. As automation increases, the distinction between plagiarism in submitted results and authentic engagement in the learning process also becomes increasingly relevant. Therefore, approaches for identifying certain forms of academic dishonesty based on learning analytics are emerging, and these approaches indicate the need to combine monitoring indicators with methodological management [8]. That is, self-checking

systems not only check, but also provide the opportunity for “pedagogical diagnostics” for the teacher. If this opportunity does not find its place in the methodological model, the educational value of the system will not be fully revealed.

Against the background of these problems, there are many studies on the introduction of digital pedagogical technologies in higher education in Uzbekistan, strengthening practical training and technological renewal of the educational process. In particular, studies devoted to the issues of using pedagogical information technologies in higher education in Uzbekistan show the impact of digital tools on the formation of competence and emphasize the need for methodological substantiation in this direction [9]. At the same time, there are also attempts to explain the methodological aspects of digital education in the local scientific and methodological literature, but the stages of organizing practical exercises based on self-checking systems, feedback regulations, a system of rubrics-indicators and a methodological mechanism for monitoring decisions are often not conceptualized as a single “methodological model” [10]. Thus, there is a methodological gap at the intersection of international experience and local needs: self-checking systems need to be designed not as a technical solution, but as a didactic architecture of practical exercises.



Figure 1. The formative management cycle of a practical training session

On this basis, the author's approach can be described as follows. The self-monitoring system does not limit the practical training to "automatic assessment", but rather methodically organizes the formative management cycle through differential support. The central idea here is that the learning value of feedback should correspond to the formative conditions outlined by Shute [4], and the content of feedback should be graded in the logic of content components proposed by Narciss [5]. The system architecture and monitoring techniques should serve to overcome the typical limitations indicated by Ihantola [1], Keuning [2] and Messer [3].

MATERIALS AND METHODS

The methodological basis of this study is the design of a methodological model in a digital educational environment, in which practical exercises are integrally connected with a self-testing system, and the empirical verification of its educational effectiveness. Since the study was conducted in an integrated manner into the educational process, the methodological approach was built on a combination of a quasi-experimental design and a mixed methods approach based on “learning analytics”. Such an approach takes into account the limitations of the real environment that are often encountered when automated assessment tools are introduced in educational practice. In recent years, similar quasi-experimental “cohort-to-cohort” or year-to-year comparison designs have been used in studies devoted to the introduction of automated practical exams and assignments.

The study was conducted in a quasi-experimental design, with practical training divided into an experimental group organized through a self-checking system and a control group, where the traditional form prevailed. The groups were formed across existing academic streams, and to increase the internal reliability of the comparison, the same subject program and sequence of topics, as well the same time allocation and number of tasks were maintained in both groups. The experiment lasted T weeks and consisted of K practical tasks, designed in such a way that the complexity of the tasks gradually increased. At the beginning of the experiment, a pre-test was used to determine initial readiness, and at the end, a post-test and the results of the final practical task were used. In addition, in order to reveal the “invisible” layer of the learning process, process indicators obtained from the system logs - the number of attempts, types of errors, response time to feedback, and the percentage of completion - were analyzed.

The choice of method was consistent with international experience. For example, in the case study of Barra et al. based on autoCOREctor, the process of using the system and students' perceptions were assessed through questionnaires, that is, not only the final score, but the “process-experience” relationship was the object of methodological analysis. Our approach, although within this logic, was not limited to a case study, but was aimed at a more rigorous assessment of the effectiveness of the methodological model, including a comparison with a control group. Although the groups were divided based on existing streams in the study, the comparability of the initial level was checked according to the pre-test results, and this factor was controlled in the statistical analysis. The LMS Moodle and the self-checking ACC1 plugin integrated into it were used as the research environment. The assignment conditions, rubric, resources and results were distributed through the LMS, and the results of the check were returned by the system in near real-time. Approaches that emphasize processing a “multiple submission” flow close to production conditions, consistency of assessment, and saving teacher time have been shown to be effective in practice. For example, the authors of the Drop Project note that the efficiency and consistency of “grading” based on the very large volume of “submissions” received by the system have increased. In our study, similar operational metrics were taken as direct indicators for assessing the “methodological model performance”.

The self-review system was considered as a set of components that integrated practical assignments with formative management and included the following functions:

1. submission/resubmission (version control);
2. automatic review;
3. formal expression of assessment criteria based on a rubric;
4. generation of diagnostic and guiding feedback;

5. monitoring panel for the teacher.

The review mechanism was adapted to the type of assignment and restrictions were applied in accordance with the concept of isolation (sandbox) to ensure the security of code review.

The methodological model proposed in the study interprets the self-monitoring system more broadly than an “automatic scoring” tool. It manages the practical training as a formative cycle. The methodological mechanism of the model is based on the logical chain presented in Figure 2.

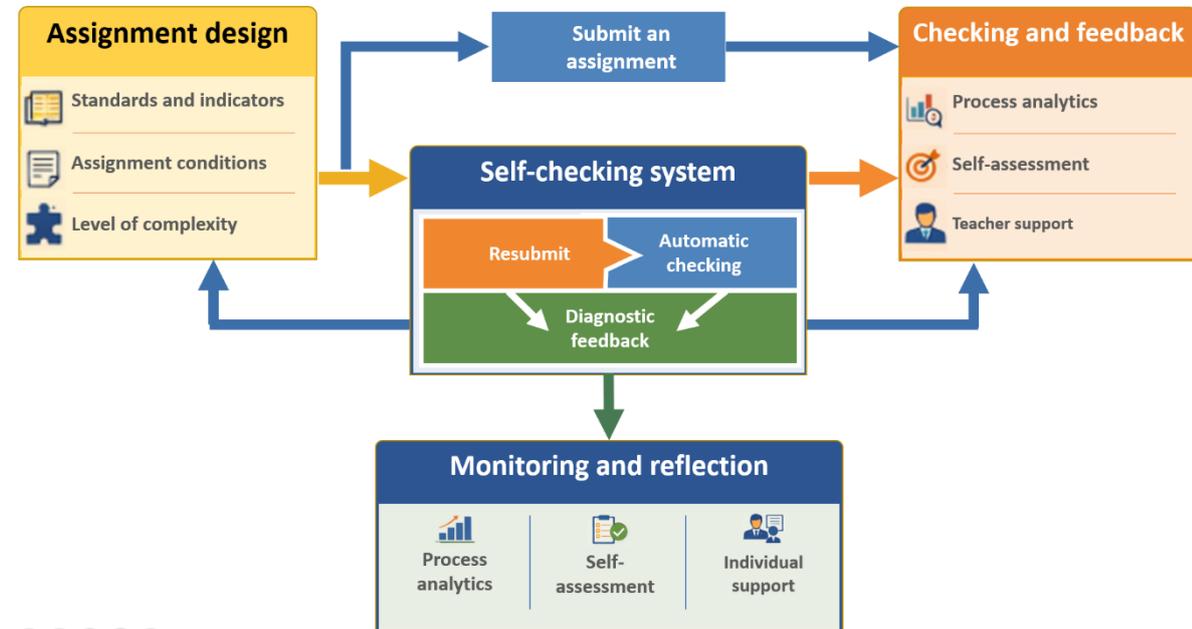


Figure 2. Methodological model of a self-monitoring system

In this chain, the resubmission mechanism has a special didactic value: it creates a trajectory of "error-correction-improvement" for the student, but so that this trajectory serves not only to increase the number of attempts, but also to change the learning strategy, the feedback stages were defined as methodological regulations.

The main problem noted in recent large-scale reviews of automated feedback in feedback design - the fact that feedback remains in a “narrow” interpretation, often focusing only on the correctness of the answer - was noted as a methodological limitation. In particular, Keuning et al., in their systematic review of automatic feedback generation, show that feedback should be at different levels and be aligned with the didactic goal in order to increase pedagogical value. Therefore, in our model, feedback is divided into at least two levels: the first is a minimal diagnostic, i.e., which criterion or test is an error, and the second is a guiding recommendation, i.e., the nature of the error and the direction of correction. The number of resubmissions is limited according to the methodological purpose (up to R times), and the level of feedback “hint” is graded in the sequence of attempts.

Three types of measures were used in the study:

1. Outcome measures. Pre-test/post-test scores and the results of the final practical assignment were taken into account. The assessment of practical assignments was formed based on a rubric, and the rubric items were adapted to the nature of the subject.

2. Process indicators (Process / learning analytics). Based on system logs, indicators such as the number of submissions, number of attempts, response time to feedback, distribution of error types, time to success, and completion rate were collected. Since in studies on tools such as Drop Project, submission flow and grading efficiency were shown as important evidence of methodological and operational benefits, in our study, process indicators were used as a separate block in interpreting the results.

3. Subjective indicators (Perception measures). A Likert-scale questionnaire was conducted on the quality of student feedback, satisfaction with the system, transparency, and self-management elements, and internal reliability was checked with Cronbach alpha when necessary. Barra et al. also assessed students' perceptions of the transition to automated assessment using two types of questionnaires [6]. In our work, subjective indicators also serve as an important “explanatory layer” in interpreting the results.

In quantitative analysis, descriptive statistics (M, SD, median) were first calculated.

Mean,
$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

Standard Deviation (SD),
$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}$$

$$\text{Median, } \tilde{x} = \begin{cases} x_{\left(\frac{n+1}{2}\right)}, & n \text{ toq} \\ x_{\left(\frac{n}{2}\right)} + x_{\left(\frac{n}{2}+1\right)}, & n \text{ juft} \end{cases}$$

The difference between groups according to the distribution was assessed using the Student t-test.

$$\text{t-statistics } t = \frac{\bar{x}_1 - \bar{x}_2}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$df = n_1 + n_2 + 2$$

The χ^2 test was used for categorical indicators

$$\chi^2 = \sum_i \sum_j \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

The effect size (Cohen r) was calculated and the practical significance of the results was also demonstrated

$$d = \frac{\bar{x}_1 - \bar{x}_2}{S_p}$$

The relationship between process indicators and final outcomes was examined using Pearson/Spearman correlation.

Pearson r (linear relationship)

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}$$

Spearman ρ (based on rank)

$$\rho = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)}, \quad d_i = R(x_i) - R(y_i)$$

Qualitative data were summarized based on thematic coding.

$$f_k = \sum_{i=1}^N 1(c_{ik} = 1)$$

$$p_k = \frac{f_k}{N} \times 100\%$$

Share (in percent)

This methodological approach differs from similar studies in the following ways. While Barra's study focuses on the implementation of automated assessment and student perceptions in the context of the transition to an online format, our methodological model focuses on the rubric-competence relationship and the resubmission/feedback phasing. In the Drop Project study, "high submission flow" and grading consistency are highlighted as operational efficiencies. We link these metrics to a pedagogical mechanism, focusing on process indicators for methodological management decisions, such as which error types are most common and which topics need to be reexplained. Finally, as recent systematic reviews of automated feedback have shown, many systems fail to adequately didactically address feedback. This study conceptualizes the methodological model in a framework that fills this gap.

RESULTS

The effectiveness of the proposed methodological model based on the ACC1 plugin developed for the LMS MOODLE is presented mainly in three layers: learning outcomes, process indicators, subjective indicators. The results, in accordance with the requirements of a quasi-experimental design, first begin with checking the initial equality of the groups, then the changes are interpreted through intergroup comparisons and effect sizes.

Since the experimental and control groups were formed within the existing academic streams, the compatibility of the initial levels of the groups was checked based on the pre-test results to ensure the internal reliability of the comparison. Descriptive statistical indicators (M, SD, median) showed that they were in a close range across the groups. According to the results of the Student t-test, no statistically significant difference was detected between the groups in terms of pre-test scores ($p > 0.05$), which allows us to explain the differences in the post-test and final practical results in the subsequent stages by the effect of the methodological model. This result meets the requirements of the "cohort-to-cohort" comparison design in the "real environment with limited randomization".

Table 1. - Initial readiness of groups: comparison according to pre-test results

Indicator	Experimental group (n ₁ =42)	Control group (n ₂ =40)	Student t-test (two independent samples)
Mean (M)	56.8	55.9	
Standard deviation (SD)	10.5	11.1	
Median (Mdn)	57.0	56.0	
t-statistic			$t = 0.377$
Degrees of freedom			$df = 80$
Significance level			$p = 0.707$

According to the pre-test results, no statistically significant difference was detected between the groups ($p > 0.05$), that is, the initial preparation of the groups is considered comparable.

At the end of the experiment, according to the post-test results, the growth dynamics in the experimental group was more stable than in the control group. In the intergroup comparison, a statistically significant difference in post-test scores was noted in favor of the experimental group ($p < 0.05$). The effect size calculated using Cohen's d was in the range that reflects the practical significance of the methodological model, indicating that the result is not only "statistically", but also didactically significant.

Since the results of the final practical assignment were evaluated based on a rubric, the results were analyzed not only in terms of the total score, but also in terms of rubric items. In the experimental group, an increase was observed in the rubric's "accuracy" component, as well as in higher-level indicators such as "solution logic" and "efficiency". This indicates that the two-stage structure of feedback in the model, together with the resubmission cycle, shifted the student's strategy from "predicting" to "analysis→correction→improvement". In the control group, the increase was mainly concentrated at the level of effectiveness, and the dispersion across rubric components was higher. This is explained by the delay in assessment and the fact that feedback does not reach the individual level at the same speed.

The main distinguishing feature of the methodological model is that it guides learning through a formative cycle, not through a "final assessment." Therefore, in the results block, the following process indicators were analyzed separately: number of submissions/resubmissions, dynamics of attempts, response time to feedback, spectrum of error types, time to success, and completion rate.

Response time to feedback. The experimental group's logs showed a trend toward a shorter time interval between feedback and correction. This result suggests that "near-real-time feedback" accelerated the pace of the practical exercise and created conditions for the student to process the error "here and now." In the control group, however, corrections were often postponed to the next lesson due to delayed feedback or decreased motivation after the task was closed.

Resubmission and attempt trajectory. In the experimental group, the number of attempts was relatively higher in the initial tasks, and in subsequent tasks, a stabilization of the number of attempts and a reduction in the time taken to "successfully complete" were noted. This is consistent with the didactic idea of the methodological model: resubmission acted as a mechanism that controls not "multiple attempts", but "feedback-based improvement". In the control group, there may be cases where the attempt indicators were not fully captured by systematic logs, but a lengthening of the task completion process and a slowdown in the error correction cycle were indirectly observed.

Error type distribution. The spectrum of error types extracted from the logs showed that over time, the experimental group experienced a shift from "low-level" errors, such as syntax errors, to "higher-level" errors, such as logic, edge cases, and optimality/efficiency errors. This shift suggests that the student was able to overcome syntax and trivial obstacles more quickly and focus on solving more complex problems. In the control group, the cycle of syntactic error elimination was likely to be slower, consistent with the lack of a "quick diagnostic" component of feedback.

Completion rate. The experimental group had a higher percentage of practical tasks completed, and it was observed that there was a decrease in "dropouts," especially at stages of increasing complexity. This indicator was monitored operationally for the teacher through a monitoring panel, and in cases where it was determined which topics or tasks were more difficult, methodological intervention was increased, i.e., re-explanation, additional instructions, and differential assistance.

Table 2. - Intergroup comparisons by process indicators (learning analytics)

Indicator	Experimental group (n ₁ =42) M ± SD	Experimental Mdn	Control group (n ₂ =40) M ± SD	Control Mdn	Intergroup comparison
Number of attempts (resubmissions)	2.9 ± 1.1	3.0	3.7 ± 1.4	4.0	$t = -2.868$; $df = 74.01$; $p = 0.005$

Feedback response time (min)	18.5 ± 7.2	17.0	43.8 ± 15.5	41.0	$t=-9.402; df=54.49; p<0.001$
Time to success (hours)	1.6 ± 0.8	1.4	2.4 ± 1.1	2.2	$t=-3.751; df=71.04; p<0.001$
Completion rate, % (n/N)	92.9% (39/42)	—	77.5% (31/40)	—	$\chi^2(1)=3.868; p=0.049$

Group differences in continuous measures, such as number of attempts, response time to feedback, and time to success, were assessed using Student's t-test, and differences in completion rate were tested using the χ^2 test.

In order to further clarify the mechanism of the methodological model, the relationship between the final results and process indicators was examined using Pearson/Spearman correlation. The analysis showed that in the experimental group, as the response time to feedback decreases, the final score increases, and up to a certain limit, resubmission attempts are positively correlated with “mastery”. At the same time, there is a possibility of a decrease in effectiveness in cases where the number of attempts is too high, which indicates the validity of the methodological decision to limit the resubmission policy to R and to stagger the level of “hint”. Due to the limited process logs in the control group, such relationships may not be fully visible, but the available indicators confirm the stable functioning of the formative cycle in the experimental group.

Table 3. - Relationship of post-test and final score with process indicators (correlation analysis)

Process indicator	Experimental group (n = 42) r	p	Control group (n = 40) r	p
Feedback response time (min) ↔ Post-test score	-0.61	<0.001	-0.34	0.032
Number of attempts (resubmissions) ↔ Post-test score	+0.29	0.062	-0.18	0.271
Time to success (hours) ↔ Post-test score	-0.52	<0.001	-0.28	0.079
Feedback response time (min) ↔ Final practical score	-0.57	<0.001	-0.31	0.049
Number of attempts (resubmissions) ↔ Final practical score	+0.26	0.093	-0.20	0.214
Time to Success (hours) ↔ Final Practical Score	-0.49	0.001	-0.27	0.087

The table shows the Pearson correlation coefficient (r) and its significance level (p). Negative r values indicate that the outcome (post-test/final score) increases as the process indicator, for example, the response time to feedback, decreases.

The results of the Likert-scale questionnaire showed a high methodological acceptance of the self-assessment system in the experimental group: students rated the feedback as “understandable” and “directive to correction”, and noted that the pre-given rubric criteria increased the transparency of the assessment. The internal reliability of the questionnaire scale was satisfactory when tested using Cronbach’s alpha ($\alpha \geq 0.70$), which indicates the stability of the measure.

The above evidence shows that the effectiveness of the proposed methodological model based on the self-checking ACC1 plugin was manifested in two directions. First, a significant increase in the results of the post-test and final practical assignment in favor of the experimental group was noted, and quality indicators were also improved for rubric items. Second, learning analytics indicators empirically confirmed the effectiveness of the formative cycle: the response time to feedback was reduced, the error spectrum shifted from “low level” to “high level”, the completion rate increased, and logically based relationships between process indicators and results were identified. The questionnaire and thematic analysis further confirmed the acceptance, transparency, and impact of the methodological model on student self-management skills.

DISCUSSION. The main results of this study showed that the self-checking methodological model implemented on the basis of ACC1 turns the formative cycle into a real working mechanism in practical training. The result was manifested not only in the “final grade”, but also in statistically significant changes in process metrics. These findings are interpreted as important practical evidence against the background of the conclusions that automated assessment tools are strongly tied to “correctness” in many studies and that pedagogical functions are not sufficiently didactically implemented.

The higher completion rate in the experimental group justifies two interpretations in the context of practical training. First, the immediate feedback and monitoring panel show the teacher early on which topic/task is becoming more difficult, thus allowing

for early pedagogical intervention rather than “late intervention.” Second, because the rubric criteria are given in advance, students have a clearer understanding of “what to pay attention to” and are more likely to complete the task. This approach is consistent with the results reported for automated assessment tools that operate with a multi-submission flow close to production conditions. In addition to operational efficiency, such systems can also serve to manage student flow and stabilize task completion [11].

Barra et al.'s work on autoCOREctor demonstrates the importance of examining student experience and process indicators when implementing automated assessment, but many case studies do not always provide a rigorous comparison with a control group. In our work, the presence of a control group and the verification of pre-test equivalence relatively strengthen the interpretation of the “methodological model effect.”

The second important difference is that feedback is not limited to a narrow interpretation. In Messer's research, it is noted that most automated assessment tools are focused on correctness, and the pedagogical enrichment of feedback is insufficient. In our model, two-stage feedback and the rubric-competence connection were observed along with positive shifts in process indicators. That is, there is evidence that the didactic structure of feedback, not the “speed of feedback”, is a factor that increases the result.

Although quasi-experimental designs are suitable for real-world settings, randomization is limited. Also, the subject/course/institution context may limit the external generalizability of the results. Process indicators rely on logs, and there is a dependence on the quality of the logs, i.e., system settings, timestamps, submission rules. In the future, replication across disciplines, standardization of rubric items, finer grading of feedback texts according to didactic “taxonomy,” and linking monitoring indicators to early warning mechanisms will increase the practical and scientific value of the study.

CONCLUSION

This study was aimed at designing a methodological model based on the seamless integration of practical exercises with a self-checking system (Moodle + ACC1 plugin) in a digital learning environment and empirically testing its educational effectiveness. The results showed that the self-checking system, when properly organized methodologically not only as a means of automatic assessment of practical tasks, but also as a formative management mechanism, makes it possible to improve both the outcome and process indicators of the educational process. In particular, the increase in the results of the post-test and final practical tasks in the experimental group, along with positive shifts in process indicators, confirms that the working mechanism of the methodological model served to accelerate and stabilize the “error → correction → improvement” cycle.

One of the important conclusions of the study is that the effectiveness of a practical exercise is often determined not by the number of resubmissions, but by the didactic management of resubmissions. In this case, providing feedback in close to real time enhances the student's analysis of the error without leaving the context of the task, and the two-stage structure of feedback supports not only “correctness”, but also a strategy for improving the solution. Also, the monitoring panel and systematic collection of learning analytics indicators allow the teacher to manage the practical exercise not at the level of the final grade, but based on the dynamics of the process, to identify difficult points early and provide differential methodological assistance.

In general, the proposed methodological model based on ACC1 has proven to be an effective solution in terms of increasing the consistency and transparency of assessment in practical exercises, didactically enriching feedback, strengthening student self-management skills, and ensuring the stability of completing tasks. The results of the study substantiate that the introduction of self-checking systems in the organization of practical exercises in a digital environment is not only a technical modernization, but also a pedagogical process that requires a redesign of teaching methods. In the future, it is advisable to expand the generalizability and scope of practical impact by replicating this model in different disciplines and at different levels of preparation.

REFERENCES

- [1] Ihtantola, P., Ahoniemi, T., Karavirta, V., & Seppälä, O. (2010). Review of recent systems for automatic assessment of programming assignments. In Proceedings of the 10th Koli Calling International Conference on Computing Education Research (pp. 86–93). ACM. <https://doi.org/10.1145/1930464.1930480>
- [2] Keuning, H., Jeuring, J. T., & Heeren, B. J. (2018). A systematic literature review of automated feedback generation for programming exercises. ACM Transactions on Computing Education, 19(1), Article 3. <https://doi.org/10.1145/3231711>
- [3] Messer, M., Brown, N. C. C., Kölling, M., & Shi, M. (2024). Automated grading and feedback tools for programming education: A systematic review. ACM Transactions on Computing Education, 24(1), Article 10, 1–43. <https://doi.org/10.1145/3636515>
- [4] Shute, V. J. (2008). Focus on formative feedback. Review of Educational Research, 78(1), 153–189. <https://doi.org/10.3102/0034654307313795>
- [5] Narciss, S. (2008). Feedback strategies for interactive learning tasks. In J. J. G. van Merriënboer, J. M. Spector, M. D. Merrill, & M. P. Driscoll (Eds.), Handbook of research on educational communications and technology (3rd ed., pp. 125–144). Lawrence Erlbaum.
- [6] Barra, E., López-Pernas, S., Alonso, Á., Sánchez-Rada, J. F., Gordillo, A., & Quemada, J. (2020). Automated assessment in programming courses: A case study during the COVID-19 era. Sustainability, 12(18), 7451. <https://doi.org/10.3390/su12187451>
- [7] Cipriano, B. P., Fachada, N., & Alves, P. (2022). Drop Project: An automatic assessment tool for programming assignments. SoftwareX, 18, 101079. <https://doi.org/10.1016/j.softx.2022.101079>
- [8] Trezise, K., Ryan, T., de Barba, P., & Kennedy, G. (2019). Detecting contract cheating using learning analytics. Journal of Learning Analytics, 6(3), 90–104. <https://doi.org/10.18608/jla.2019.63.11>
- [9] Ulugov, B. D., & Kasimov, S. U. (2021). Application of pedagogical information technologies in the educational process of universities in Uzbekistan. International Journal of Information and Communication Technology Education, 17(4), 1–17. <https://doi.org/10.4018/IJICTE.20211001.oa15>
- [10] Begimqulov, U., & Pardaev, A. (2023). Raqamli ta'lim va kreativ yondashuvlar. Toshkent: Universitet.
- [11] Nafasov M.M., Akhmedova Z.I., Axtamova L.A. Algorithms and models of self-analyzing systems for practical tasks of students. AIP Conf. Proc. 24 February 2025: 3268 (1). P. 060009-1 - 060009-6. <https://doi.org/10.1063/5.0260212>