

Using TIMSS International Assessment to Develop Natural Science Literacy in Prospective Chemistry Teachers

Gulnoza Asadullayeva Asadulla Qizi

Termiz State Pedagogical Institute, Termez, Uzbekistan

Corresponding Email: gulnoza.asadullayeva41@gmail.com

Abstract. This study addresses the problem of preparing prospective chemistry teachers for the cognitive requirements of TIMSS and, through this preparation, developing their natural science literacy. The objective was to design and empirically test a methodology for developing natural science literacy in prospective chemistry teachers based on TIMSS-type tasks. A quasi-experimental design with control and experimental groups was applied at three Uzbek higher education institutions (Navoiy State University, Urganch State Pedagogical Institute and Termez State Pedagogical Institute) during 2023-2026, involving 434 second- and third-year chemistry-major students. Experimental groups studied Inorganic Chemistry, Organic Chemistry and Methods of Teaching Chemistry through graphic-design, context-based tasks aligned with the TIMSS cognitive domains of knowing, applying and reasoning, while control groups received traditional instruction. Reliability of the differences was verified using Pearson's chi-square criterion. Observed values (8.72-9.36, mean 9.13) exceeded the critical value (7.815, $p=0.05$, $df=2$) at all institutions, confirming statistical significance. Experimental groups achieved 16.8% higher overall performance than control groups. It is concluded that systematic integration of TIMSS-type context tasks into chemistry teacher training curricula effectively develops natural science literacy, and corresponding recommendations for curriculum design and teacher-training modules are proposed.

Keywords: chemistry teacher training; chi-square criterion; context-based tasks; natural science literacy; pedagogical experiment; TIMSS

1 Introduction

Global transformations in social, economic and educational life have placed natural science literacy at the centre of policies aimed at preparing intellectually capable, creative specialists able to compete internationally. International large-scale assessments, most notably TIMSS and PISA, conducted under the Organisation for Economic Co-operation and Development (OECD) and the International Association for the Evaluation of Educational Achievement (IEA), have become the principal instruments through which national education systems benchmark the science and mathematics literacy of

their students against international standards. A growing body of comparative research has shown that raising national performance in these assessments depends not only on curriculum reform but, decisively, on the readiness of teachers to organise instruction around the cognitive domains that these assessments measure — knowing, applying and reasoning.

The Republic of Uzbekistan has, in recent years, undertaken a series of legislative and strategic measures intended to raise the quality of education and to align the competencies of pedagogical personnel with international requirements, including the Development Strategy of New Uzbekistan for 2022-2026 and a number of decisions concerning the organisation of international assessment studies in general secondary education. These measures presuppose that future subject teachers, including chemistry teachers, must themselves be capable of designing, analysing and applying TIMSS-type tasks before they can prepare school pupils for such assessments.

International literature confirms this direction. Studies of teacher-education reform in high-performing systems describe how programmes were reorganised around research-based, cognitively structured pedagogy to support high and stable science-assessment results [4]. Foundational technical reports on TIMSS assessment frameworks specify the cognitive domains — knowing, applying and reasoning — against which item difficulty and content coverage must be calibrated [1]. Within chemistry education specifically, recent classroom-based work has shown that graphic organisers and visually structured tasks used by and with prospective chemistry teachers can meaningfully improve conceptual understanding of demanding topics such as electrochemistry [2], while systematic reviews of inquiry-based chemistry education confirm that cognitively structured, hands-on tasks consistently produce stronger learning outcomes than lecture-based instruction alone [5]. A parallel systematic review of inquiry-based science education specifically within pre-service teacher education shows that most existing studies address the procedural, conceptual, pedagogical and affective cognitive domains of inquiry, while comparatively few examine the epistemic and social domains or the transition from learning science through inquiry to teaching science through inquiry [4]. Within the post-Soviet and Central Asian context, a substantial body of methodological work has examined the comparability of national and international quality-assessment systems [14], [15], yet comparatively little of this work has translated TIMSS cognitive requirements into a concrete, empirically tested methodology for training prospective chemistry teachers specifically.

Natural science literacy, understood as the capacity to engage with science-related issues and with the ideas of science as a reflective citizen, is composed of interrelated cognitive components: knowing established scientific facts, concepts and procedures; applying this knowledge to solve familiar and unfamiliar problems; and reasoning through analysis, synthesis, evaluation and justification in complex or novel situations [1]. A recent systematic review of two decades of TIMSS- and PISA-based research confirms that instructional features aligned with these cognitive domains consistently predict stronger science-teaching and science-learning outcomes across national contexts [3], while a large cross-national analysis of TIMSS 2019 data has shown that student- and teacher-level instructional variables remain significant predictors of science literacy even after controlling for socioeconomic and country-level factors [6]. A recent

bibliometric and content analysis of scientific-literacy research further indicates a sustained international rise in studies devoted to measuring and developing this construct, underlining its continuing relevance as a policy priority [13]. Comparative studies of national quality-assessment systems in the post-Soviet space have stressed the need to harmonise local pedagogical practice with the cognitive requirements of international benchmarks [14], [15], yet these studies have generally remained descriptive or comparative in character rather than proposing and testing a concrete instructional methodology.

For chemistry teacher education specifically, this creates a distinctive challenge. Chemistry combines abstract theoretical content (atomic and molecular structure, reaction mechanisms, thermodynamics) with practical, experimentally verifiable phenomena, and TIMSS-type context tasks in chemistry typically require students to move fluidly between symbolic, graphical and verbal representations of the same phenomenon. Recent work on scientific-practices assessment in the general chemistry laboratory confirms that carefully designed, cognitively structured tasks can reliably measure and develop exactly this kind of applied and reasoning-level competence in chemistry students [9], while studies of practising chemistry teachers' context-based teaching experiences show that teachers themselves need sustained methodological support to design and apply such tasks confidently in the classroom [10]. A quasi-experimental study on socio-scientific-issue-based inquiry learning likewise found significantly higher scientific habits of mind and chemical literacy in experimental groups than in traditionally taught control groups, reinforcing the value of structured, context-embedded instructional designs of the kind adopted here [11]. Prospective chemistry teachers must therefore not only master disciplinary content but also learn to design, adapt and sequence graphic-design and context-based tasks of calibrated difficulty for their future pupils. Without deliberate methodological support during initial teacher training, this dual requirement — disciplinary mastery combined with assessment-design competence — is unlikely to develop through traditional, lecture-based instruction alone.

This gap defines the problem addressed in the present study: although the importance of TIMSS-oriented teacher preparation is widely acknowledged, there is a lack of an integrated, discipline-specific methodology — combining didactic content, a pedagogical model and a system of context-based tasks — for developing natural science literacy in prospective chemistry teachers, together with empirical evidence of its effectiveness [15], [16], [17]. Accordingly, the purpose of this article is to present and empirically evaluate a methodology for developing natural science literacy in prospective chemistry teachers through systematic preparation for the TIMSS international assessment, combining graphic-design and individually differentiated context tasks with the cognitive domains of knowing, applying and reasoning. Two guiding questions structured the research: (1) can a system of TIMSS-aligned, graphic-design context tasks be integrated into core chemistry courses (Inorganic Chemistry, Organic Chemistry, Methods of Teaching Chemistry) without displacing required disciplinary content; and (2) does such integration produce a statistically verifiable improvement in natural science literacy compared with traditional instruction. The corresponding research hypothesis was that systematic, cognitively structured integration of TIMSS-type tasks into

subject-matter courses would yield significantly higher achievement in experimental groups than in control groups taught by conventional methods.

2 Methods

A quasi-experimental research design with parallel control and experimental groups was employed to evaluate the proposed methodology. The independent variable was the instructional methodology used in the Inorganic Chemistry, Organic Chemistry and Methods of Teaching Chemistry courses: control groups were taught by traditional, predominantly descriptive methods, while experimental groups studied the same subject content through a system of TIMSS-oriented, context-based and graphic-design tasks integrating natural science content with pedagogical and information technologies. The dependent variable was the level of natural science literacy, operationalised through differentiated achievement categories (excellent, good, satisfactory, unsatisfactory) assigned according to criteria developed specifically for the experiment.

The pedagogical experiment was carried out during the 2023-2026 academic years at three higher education institutions of the Republic of Uzbekistan: Navoiy State University, Urganch State Pedagogical Institute and Termez State Pedagogical Institute. A total of 434 second- and third-year students majoring in chemistry education participated as respondents, distributed across the three institutions as follows: 203 students at Navoiy State University (100 in control groups, 103 in experimental groups), 56 students at Urganch State Pedagogical Institute (28 in control groups, 28 in experimental groups), and 175 students at Termez State Pedagogical Institute (86 in control groups, 89 in experimental groups).

The instructional materials used in the experimental groups included TIMSS-format tasks, context-based questions, tests, presentation slides and study guides developed by the research team specifically for the development of natural science literacy, structured to correspond to the three TIMSS cognitive domains: knowing (recall of facts, concepts and procedures), applying (use of knowledge to solve problems in familiar situations) and reasoning (analysis, synthesis and justification in unfamiliar or complex situations). Assessment instruments and observation protocols were used to record student performance at the end of each academic year.

The underlying pedagogical model guiding task design specified four interlinked blocks: a target block, defining the aim of aligning chemistry instruction with TIMSS cognitive requirements; a content block, structuring disciplinary material around graphic-design tasks of individually differentiated complexity; a procedural block, prescribing adaptive and interactive instructional techniques suited to the interactive learning environment; and an evaluative block, defining the achievement criteria and the statistical procedure used to judge the methodology's effectiveness (Figure 1). Tasks were graduated in three complexity tiers corresponding to the knowing, applying and reasoning domains, and each task combined a graphic or diagrammatic element (e.g., reaction schemes, particle-level diagrams, data graphs) with a context drawn from everyday or applied chemistry, consistent with the item format used in TIMSS assessments. Students in the experimental groups worked through these tasks as an integral part of

regular seminar and laboratory sessions rather than as supplementary test-preparation drills, so that TIMSS-oriented thinking was practised within, rather than alongside, the standard curriculum.

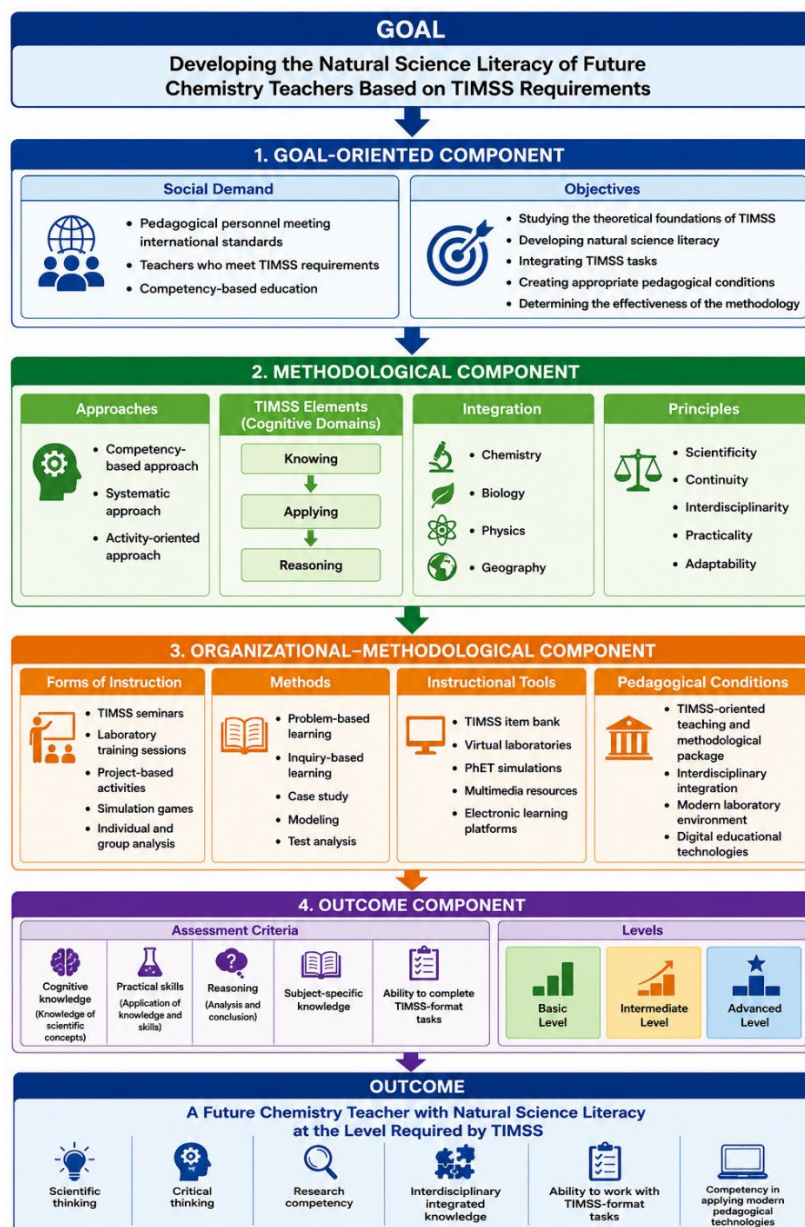


Fig. 1. Pedagogical model for developing natural science literacy in prospective chemistry teachers through TIMSS-based preparation.

The reliability of the differences between the control and experimental groups was tested using Pearson's chi-square (χ^2) criterion, appropriate for comparing the distribution of categorical outcomes (achievement levels) between two independent samples. The null hypothesis (H_0) stated that the probability of falling into each achievement category was equal for the control and experimental groups; the alternative hypothesis (H_1) stated that these probabilities differed. The test statistic was computed as

$$T_{obs} = [1 / (n_1 \cdot n_2)] \cdot \sum [(n_1 \cdot O_{2i} - n_2 \cdot O_{1i})^2 / (O_{1i} + O_{2i})],$$

where n_1 and n_2 denote the number of students in the control and experimental groups respectively, and O_{1i} and O_{2i} denote the number of students in each achievement category i in the control and experimental groups respectively. With three assessment categories used for the comparison, the degrees of freedom were $k = 2$, and at a significance level of $p = 0.05$ the critical value from the χ^2 distribution table was $T_{crit} = 7.815$. If T_{obs} exceeded T_{crit} , H_0 was rejected in favour of H_1 , indicating a statistically reliable difference between groups attributable to the experimental methodology.

3 Results and Discussion

Table 1 summarises the distribution of student achievement levels in the control and experimental groups at each of the three participating institutions, together with the computed χ^2 values and the corresponding statistical decision.

Table 1. Distribution of student achievement levels and χ^2 test results (2023-2026)

Institution / Academic years	Group	Excellent	Good	Sat- isfac- tory	Unsatis- factory	N	T_obs	Decision
Navoiy State University (2023-2026)	Control	15	38	43	4	100	9.31	H1 accepted
	Experimental	40	44	19	0	103		
Urganch State Pedagogical Institute (2023-2026)	Control	5	8	13	2	28	8.72	H1 accepted
	Experimental	18	8	2	0	28		
Termez State Pedagogical Institute (2023-2026)	Control	12	26	43	5	86	9.36	H1 accepted
	Experimental	42	30	17	0	89		
Total	214 / 220	132	154	137	11	434	9.13	H1 accepted

At all three institutions, the observed value of the χ^2 criterion exceeded the critical value of 7.815 ($df = 2$, $p = 0.05$): 9.31 at Navoiy State University, 8.72 at Urganch State

Pedagogical Institute and 9.36 at Termez State Pedagogical Institute, with an aggregate value of 9.13 across the combined sample of 434 students. In every case the null hypothesis was rejected and the alternative hypothesis accepted, confirming that the differences between the control and experimental groups were statistically reliable rather than attributable to chance. In aggregate, 132 students achieved an "excellent" rating, 154 a "good" rating, 137 a "satisfactory" rating and only 11 an "unsatisfactory" rating, with the large majority of top ratings concentrated in the experimental groups (Figure 2). Overall, students who studied under the proposed methodology reached an achievement level approximately 1.168 times, or 16.8%, higher than students taught by traditional methods.

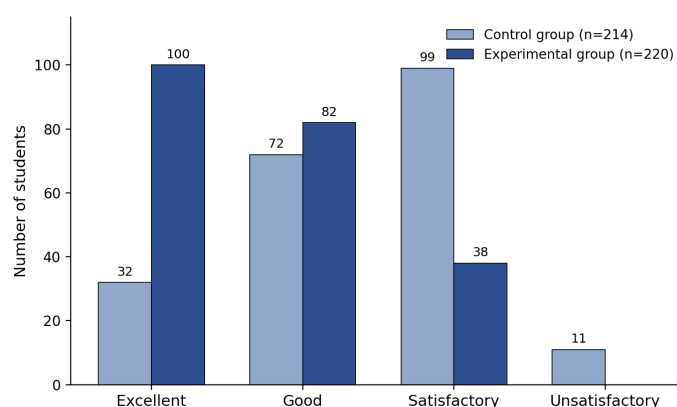


Fig. 2. Distribution of student achievement levels in control and experimental groups (aggregate, N = 434).

These findings indicate that systematically embedding TIMSS-type context tasks into subject-matter courses — rather than treating international-assessment preparation as a separate activity — measurably strengthens prospective chemistry teachers' natural science literacy. The consistency of the effect across three institutionally and geographically distinct settings (Navoiy, Urganch and Termez) suggests that the methodology's core components — differentiated task complexity, graphic-design formats and alignment with the TIMSS cognitive domains of knowing, applying and reasoning — are transferable rather than context-specific artefacts of a single institution, a pattern consistent with cross-national evidence that instructional variables aligned with TIMSS cognitive domains predict science literacy outcomes across diverse settings [3], [6].

The results are consistent with recent classroom research showing that graphic organisers used by and with prospective chemistry teachers improve conceptual understanding of demanding topics [2], and with systematic reviews confirming that cognitively structured, inquiry-oriented tasks outperform purely lecture-based instruction in chemistry education [5]. They also corroborate technical descriptions of the TIMSS assessment framework, which specify knowing, applying and reasoning as the cognitive domains against which item design should be calibrated [1]; validated instruments built on the same cognitive structure have likewise been shown to track literacy growth

reliably across school levels [8], and cognitive-diagnostic analyses of TIMSS response data confirm that fine-grained cognitive attributes, rather than aggregate scores alone, best capture differences in students' learning trajectories [7]. A comparable quasi-experimental study using socio-scientific-issue-based inquiry learning likewise reported significantly stronger chemical-literacy outcomes in experimental than in control groups [11], echoing the magnitude of the improvement observed here. The present study demonstrates that the TIMSS cognitive domains can productively organise not only pupil-facing assessment items but also the professional training of the teachers who will later prepare pupils for such assessments, extending laboratory-based evidence that structured, cognitively calibrated tasks can develop applied and reasoning-level competence in chemistry students specifically [9]. In this respect, the study extends prior systematic reviews of inquiry-based science-teacher education [4] and prior reports of chemistry teachers' context-based teaching experiences [10] toward a concrete, discipline-specific and empirically verified teacher-training methodology for chemistry education.

Several limitations should be noted. The sample, although drawn from three institutions, was confined to the chemistry-education programmes of one country, which may limit the generalisability of the findings to other subject areas or national contexts. The achievement categories used for the χ^2 analysis were institution-specific criterion measures rather than a standardised international instrument, and the study did not track long-term retention of natural science literacy beyond the period of the experiment. Future research should examine the durability of the observed gains after graduates enter classroom practice, and should extend the methodology to other natural science teacher-training programmes (biology, physics) to test its broader applicability.

4 Conclusions

The study set out to design and empirically test a methodology for developing natural science literacy in prospective chemistry teachers through systematic preparation for the TIMSS international assessment. The pedagogical experiment, conducted over three academic years at three higher education institutions with 434 participating students, confirmed that a methodology combining differentiated, graphic-design context tasks with the TIMSS cognitive domains of knowing, applying and reasoning produces a statistically reliable and practically meaningful improvement in student achievement, with experimental groups outperforming control groups by 16.8% and χ^2 values consistently exceeding the critical threshold across all three institutions.

On this basis, the following recommendations are proposed. First, curricula and subject programmes of chemistry-teacher-education tracks in pedagogical higher education institutions should incorporate a system of graphic-design and competency-based tasks aligned with the content of TIMSS international assessments. Second, interactive, inquiry-based and adaptive instructional technologies oriented toward the knowing, applying and reasoning components of natural science literacy should be used systematically to strengthen the professional competence of prospective chemistry teachers. Third, specialised training modules and methodological guides on constructing,

analysing and applying TIMSS-type tasks should be developed and introduced into both pre-service teacher preparation and in-service professional development programmes for chemistry teachers.

5 Acknowledgement

The author would like to thank the administration and chemistry-education students of Navoiy State University, Urganch State Pedagogical Institute and Termez State Pedagogical Institute for their participation in the pedagogical experiment and their cooperation throughout the 2023-2026 academic years.

References

1. I. V. S. Mullis and M. O. Martin, TIMSS 2019 Assessment Frameworks. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College, International Association for the Evaluation of Educational Achievement (IEA), 2019, 68 p.
2. C. Nakiboğlu and N. Nakiboğlu, "Views of prospective chemistry teachers on the use of graphic organizers supported with interactive PowerPoint presentation technology in teaching electrochemistry concepts," *International Journal of Physics & Chemistry Education*, vol. 13, no. 3, pp. 47-63, 2021. <https://doi.org/10.51724/ijpce.v13i3.216>
3. N. Teig, R. Scherer, and R. V. Olsen, "A systematic review of studies investigating science teaching and learning: over two decades of TIMSS and PISA," *International Journal of Science Education*, vol. 44, no. 12, pp. 2035-2058, 2022. <https://doi.org/10.1080/09500693.2022.2109075>
4. T. T. S. Strat, E. K. Henriksen, and K. M. Jegstad, "Inquiry-based science education in science teacher education: a systematic review," *Studies in Science Education*, vol. 60, no. 2, pp. 191-249, 2024. <https://doi.org/10.1080/03057267.2023.2207148>
5. K. M. Jegstad, "Inquiry-based chemistry education: a systematic review," *Studies in Science Education*, vol. 60, no. 2, pp. 251-313, 2024. <https://doi.org/10.1080/03057267.2023.2248436>
6. J. Marôco, H. Harju-Lukkainen, and J. Rautopuro, "Worldwide predictors of science literacy in lower-secondary students: a TIMSS 2019 analysis," *International Journal of Science Education*, pp. 580-598, 2024. <https://doi.org/10.1080/09500693.2024.2394239>
7. Z. Zhu, "International comparative study of learning trajectories based on TIMSS 2019 G4 data on cognitive diagnostic models," *Frontiers in Psychology*, vol. 14, Art. 1241656, 2023. <https://doi.org/10.3389/fpsyg.2023.1241656>
8. L. Zhang, H. Feng, and X. Liu, "Development and validation of an instrument for assessing scientific literacy from junior to senior high school," *Disciplinary and Interdisciplinary Science Education Research*, vol. 5, Art. 21, 2023. <https://doi.org/10.1186/s43031-023-00093-2>
9. N. S. Stephenson, E. M. Duffy, E. L. Day, K. Padilla, D. G. Herrington, M. M. Cooper, and J. H. Carmel, "Development and validation of scientific practices assessment tasks for the general chemistry laboratory," *Journal of Chemical Education*, vol. 97, no. 4, pp. 884-893, 2020. <https://doi.org/10.1021/acs.jchemed.9b00897>

10. F. M. Mustafaoğlu and A. S. Yücel, "Context-based teaching experiences of chemistry teachers: expectations, gains and applicability conditions," *Journal of Turkish Science Education*, vol. 19, no. 3, 2022.
11. A. K. Prodjosantoso and A. R. E. Nugraheni, "Promoting students' scientific habits of mind and chemical literacy using the context of socio-scientific issues on the inquiry learning," *Frontiers in Education*, vol. 6, Art. 660495, 2021. <https://doi.org/10.3389/educ.2021.660495>
12. D. A. Urdanivia Alarcon, F. Talavera-Mendoza, F. H. Rucano Paucar, K. S. Cayani Caceres, and R. Machaca Viza, "Science and inquiry-based teaching and learning: a systematic review," *Frontiers in Education*, vol. 8, Art. 1170487, 2023. <https://doi.org/10.3389/educ.2023.1170487>
13. N. Kutlu-Abu, K. Bozgün, and Ş. Uluçınar-Sağır, "Studies on scientific literacy in primary education: a bibliometric and content analyses," *Hungarian Educational Research Journal*, vol. 14, no. 2, pp. 158-183, 2024. <https://doi.org/10.1556/063.2023.00220>
14. M. T. Ergasheva, "Methodology of using assessment programs (PISA, TIMSS) in monitoring students' natural science literacy," Ph.D. dissertation abstract, Tashkent, Uzbekistan, 2023, 77 p.
15. I. E. Shernazarov, "Methodology for developing natural science literacy of prospective chemistry teachers based on international assessment studies," Dr.Sc. dissertation, Tashkent, Uzbekistan, 2023, 376 p.
16. I. E. Shernazarov, A. Yu. Iskandarov, and S. G. Khasanova, *PISA Context Tasks in Chemistry for General Secondary School Students*. Tashkent: 2021, 298 p.
17. I. E. Shernazarov, M. U. Tilyabov, G. A. Asadullayeva, and N. A. Yusupova, *Using International Assessment Methods in Teaching Chemistry*. Tashkent: Impress Media, 2025, 375 p.