DOI:10.59787/2413-5488-2026-54-2-200-212

[1]Kaisar Tursyngozhayev*, [1]Zhadyra Kaiyngbayeva, [2]Nusret Kavak

[1]Abai Kazakh National Pedagogical University, Almaty, Kazakhstan.
[2]Gazi University, Ankara, Turkey

DIGITAL GAME-BASED LEARNING IN SCHOOL CHEMISTRY AND MATHEMATICS: DUAL CASE STUDIES IN STOICHIOMETRY AND PROBABILITY THEORY AND IMPLICATIONS FOR HIGHER EDUCATION TEACHER PREPARATION IN KAZAKHSTAN

Abstract. This article reports on two complementary multi-year case studies of integrating open-access PhET Interactive Simulations into the teaching of two notoriously demanding topics in the Kazakhstani school curriculum – stoichiometry in chemistry and probability theory in mathematics – and translates the resulting evidence into a unified set of design principles and curricular recommendations for higher education teacher preparation. Case 1 examines the use of "Reactants, Products and Leftovers" in 8th-grade chemistry across three Almaty schools; Case 2 examines the use of "Plinko Probability" in 10th-11th-grade mathematics across two Almaty institutions. Both interventions used a structured Predict-Observe-Explain design with parallel 15-item pre/post-tests, student questionnaires, and lesson-time logs. Across both subjects, simulation-supported instruction was associated with substantial gains in conceptual understanding (median Cohen's d ≈ 1.0-1.1), increased intrinsic engagement, and a measurable reduction in the time required to reach class-level mastery. Crucially, the same four cross-cutting themes emerged in both subjects: visualisation of abstract sub-microscopic and stochastic phenomena, productive failure under low-stakes simulation, the limited appeal of competitive game elements, and the centrality of teacher coaching. We argue that this convergence across two very different STEM subjects strengthens the case for embedding digital game-based learning (DGBL) into Kazakhstani higher education through a four-component framework – conceptual orientation, simulation literacy, gamified lesson design and reflective practicum – applicable equally to chemistry and mathematics teacher preparation. Implications are discussed for university curriculum developers, methodologists and policymakers seeking to scale evidence-based gamification across the school-university continuum.

Keywords: digital game-based learning; PhET simulations; stoichiometry; probability theory; mathematics education; chemistry education; higher education; teacher preparation; gamification.

Introduction

The digitalization of higher education has become a strategic priority and is explicitly recognised as one of the core directions of the State Programme for the Development of Education and Science of the Republic of Kazakhstan. Within this agenda, science and mathematics teacher preparation occupies a particularly demanding position: future chemistry and mathematics teachers must master both the discipline and a rapidly expanding portfolio of digital pedagogies that they will be expected to deploy in classrooms upon graduation.

Two topics in the school STEM curriculum are routinely identified by Kazakhstani methodologists and international researchers alike as among the most cognitively demanding: stoichiometry in chemistry and probability theory in mathematics. Stoichiometry requires students to coordinate macroscopic, sub-microscopic and symbolic representations of matter, and to perform proportional reasoning over balanced equations. Probability theory requires students to reason under uncertainty, to coordinate empirical relative frequencies with theoretical models, and to overcome a set of robust intuitive misconceptions documented since Fischbein (1975). International studies report persistent learner difficulties in both subjects – difficulties that propagate upward into university and that pre-service teachers tend to replicate in their own classrooms. In the Kazakhstani context, Kaiyngbayeva et al. (2021) describe "the methodological unpreparedness of teachers" as the central obstacle to teaching probability and statistics in secondary schools, while parallel concerns are documented for chemistry teacher preparation by Karmanova et al. (2024).

The need to strengthen teacher preparation specifically for these two topics is rooted in documented school. Kazakhstan's performance in successive international assessments illustrates the scale of the problem: in PISA 2022, 15-year-olds scored 425 in mathematics and 423 in science, well below the OECD averages of 472 and 485, and only about half reached at least Level 2 proficiency in mathematics, compared with roughly two-thirds across the OECD (OECD, 2023). Such outcomes are concentrated also in the kinds of proportional, multi-step and probabilistic reasoning that stoichiometry and probability theory demand. At the level of everyday classroom practice, both topics are routinely taught as formula-driven, symbol-manipulation exercises: stoichiometry is reduced to mechanical mole calculations divorced from the particulate nature of matter, and probability is reduced to combinatorial formulae applied without any experimental or simulation experience. Probability and statistics is, moreover, a comparatively recent addition to the Kazakhstani school curriculum and no dedicated university course on its teaching methodology (Kaiyngbayeva et al., 2021). The result is a self-reinforcing cycle in which teachers reproduce the formal, non-conceptual instruction. The low baseline pre-test scores recorded in both of our own cases means of 6.8 out of 15 in chemistry and 5.9 out of 15 in probability (see Results) quantify this problem directly in the very classrooms studied here. It is this concrete gap in school practice the present study seeks to address by preparing future teachers to teach these specific topics differently.

Digital game-based learning (DGBL) and gamification have emerged as one of the most actively researched responses to such challenges across STEM. Three-level meta-analytic evidence shows that game-based learning has a positive overall effect on chemistry achievement and motivation across educational levels (Hu et al., 2022); systematic reviews of gamification confirm small-to-moderate but reliable effects on engagement, motivation and academic performance in science education (Kalogiannakis et al., 2021; Zainuddin et al., 2020); and reviews of probability and statistics education conclude that simulation-based learning environments are particularly well suited to building students' conceptual and modelling-oriented understanding of uncertainty (Pratt & Kazak, 2018). Within both chemistry and mathematics specifically, the open-access PhET Interactive Simulations developed at the University of Colorado Boulder have become a de facto international standard, with a substantial body of empirical work documenting gains in conceptual understanding, attitudes and self-efficacy when they are used as part of inquiry-oriented instruction (Perkins et al., 2006; Salame & Makki, 2021; Banda & Nzabahimana, 2023).

In Kazakhstan, however, much of the available evidence on PhET- and game-based learning is concentrated at the school level and is fragmented across subjects: chemistry studies and probability/statistics studies have largely developed in parallel rather than in dialogue. Higher education institutions, including those preparing future chemistry and mathematics teachers, have so far adopted such tools unevenly. Recent national studies emphasise that the professional competence of future chemistry teachers depends critically on their exposure to digital technologies during initial training (Karmanova et al., 2024); a parallel comparative study of secondary and higher education in Kazakhstan reports that gamification remains fragmented and lacks systemic policy support (Sardarova et al., 2026). Bridging these gaps requires evidence that is not only empirical but also cross-disciplinary demonstrating that the same digital-pedagogical principles can be translated into the methodology courses of multiple subject teacher programmes.

The evidence base on the digitalisation of Kazakhstani higher education, and on the digital-pedagogical preparation of future teachers in particular, has grown rapidly but remains thin. Beyond the two studies noted above, recent work points consistently to a gap between policy ambition and classroom-ready competence. Abiltayeva et al. (2025), surveying 240 pre-service biology teachers across three Kazakhstani universities, found high self-reported technological-pedagogical knowledge coexisting with markedly lower frequency of actual digital-tool use, and concluded that hands-on, subject-specific training is needed to close the gap between positive attitudes and practice. Yespenbetova et al. (2024) report that online and virtual laboratory work is becoming central to undergraduate chemistry but is adopted unevenly across institutions, while Sardarova et al. (2026) document an urban–regional digital divide and the absence of explicit institutional policies on gamification. Across these studies a common pattern emerges: national strategy (including the "Digital Kazakhstan" programme and Bologna-aligned reforms) endorses digital pedagogy, yet pre-service teachers graduate without structured, discipline-specific experience of designing and enacting technology-rich lessons. What is still missing is cross-disciplinary classroom evidence showing that a single set of digital-pedagogical principles can be translated into the methodology courses of more than one subject.

The present article addresses that gap by reporting on two complementary case studies conducted by the authors. Case 1 examines the integration of the PhET simulation "Reactants, Products and Leftovers" into 8th-grade chemistry instruction across three Almaty schools. Case 2 examines the integration of the PhET simulation "Plinko Probability" into 10th-11th-grade mathematics instruction (probability theory and statistics) across two Almaty schools. The article pursues four objectives: 1) to evaluate the impact of structured PhET-based interventions on students' conceptual understanding, motivation and instructional time in two distinct STEM subjects; 2) to identify the design principles that made the interventions effective at the school level; 3) to test whether the same principles emerge across chemistry and mathematics; and 4) to translate the principles into a unified framework for embedding DGBL into university methodology courses for future chemistry and mathematics teachers in Kazakhstan.

In doing so, the article contributes to three focus areas, methodology of teaching in higher education, quality assurance, and digitalization of higher education, by showing how cross-disciplinary research evidence generated in schools can be systematically used to upgrade the digital-pedagogical preparation of university students who will themselves become subject teachers.

DGBL and educational simulations are most often grounded in three converging theoretical perspectives, all of which apply equally to the teaching of chemistry stoichiometry and to the teaching of probability theory. Constructivist learning theory holds that learners actively build mental models through interaction with their environment, and predicts that interactive simulations which allow manipulation of variables and immediate visual feedback should support deeper conceptual restructuring than passive instruction (Hammad et al., 2020). Cognitive load theory complements this view by arguing that well-designed simulations can reduce extraneous load (e.g., the symbolic complexity of formulas, or the abstractness of theoretical probabilities) by externalising representations, freeing working memory for schema construction; conversely, poorly designed games may add load and depress learning (Hawlitschek & Joeckel, 2017). Self-determination theory frames the motivational benefits of gamification: feedback, autonomy of pace, and graded challenge support intrinsic motivation, particularly when game elements such as points and leaderboards are balanced against opportunities for collaboration and meaningful choice (Zainuddin et al., 2020).

Together, these frameworks predict that digital simulations – whether of molecular reactions or of stochastic experiments – should be most effective when integrated into instruction that is conceptually focused, scaffolded by structured tasks, and accompanied by metacognitive reflection. They also predict that simply providing students with a simulation, without instructional guidance, will produce small or null effects – a prediction confirmed by experimental work on game design (Hawlitschek & Joeckel, 2017). The teacher's role in shaping the simulation-based learning environment emerges as a key finding in the probability literature as well (Pratt & Kazak, 2018), suggesting that the cross-subject convergence we report below is not coincidental.

PhET Interactive Simulations are research-based, openly licensed simulations covering most of the core topics in school and undergraduate chemistry and mathematics. In chemistry, Salame and Makki (2021), in a study of 158 General Chemistry II students at the City College of New York, reported that students perceived PhET simulations as significantly enhancing their conceptual grasp of abstract topics. Banda and Nzabahimana (2023), using a quasi-experimental design with 280 secondary students in Malawi, found large effects of PhET-based instruction on academic achievement (Cohen's $d = 1.14$) along with significant gains in self-efficacy and active learning strategies. In mathematics, the "Plinko Probability" simulation – a digital implementation of the Galton board – has been widely adopted to teach binomial and normal distributions, the law of large numbers, expected value and variance through dynamic visualisation: students drop balls through a peg array and observe how empirical histograms converge on theoretical distributions as the sample size grows. Pratt and Kazak's (2018) review of probability research highlights such simulation-based environments as one of the most promising routes to addressing the persistent intuitive misconceptions documented by Fischbein (1975), and the GAISE framework explicitly endorses simulation as a core method for K-12 statistics instruction (Franklin et al., 2007).

Beyond PhET specifically, meta-analytic evidence shows that game-based chemistry learning yields a positive overall effect on student outcomes (Hu et al., 2022); a recent systematic review of gamification in primary and secondary education identified increased motivation and engagement as the most consistently reported outcomes (Vrcelj et al., 2023). At the higher education level, Montenegro-Rueda et al. (2023) report that gamified university courses are associated with stronger engagement and modest gains in achievement, while Park and Kim (2021) show that gamified online learning sustains motivation in distance settings. A broad review of computer-based technology and student engagement concludes that interactive technologies are among the most consistent predictors of behavioural and cognitive engagement when integrated thoughtfully (Schindler et al., 2017). Stohl (2005) makes a complementary argument specifically for probability teacher education: pre-service teachers who do not themselves experience probability as an experimental, simulation-rich subject tend to teach it as a purely formal topic, perpetuating their own misconceptions.

A more recent and still emerging strand of research focuses specifically on how gamification and DGBL should be taught to future teachers. Jiménez-Valverde et al. (2024), in a 14-week mixed-methods study of 65 pre-service primary teachers, demonstrated that structural gamification of a science methodology course significantly improved pre-service teachers' attitudes toward physics and chemistry and their motivation to teach these subjects. Lampropoulos and Kinshuk (2024), in a systematic review, argue that integrating gamification with virtual and immersive technologies is becoming a defining feature of next-generation teacher education. These findings suggest a dual mechanism: experiencing gamification as a learner enables future teachers to internalise its pedagogical logic, while subsequent design tasks consolidate the corresponding professional competence. Without this dual exposure, university graduates often default to the lecture-and-drill style that they themselves experienced in school, where new teachers "cannot conduct this subject in the same way as it was taught to us" because, in many cases, it was not taught to them at all.

In Kazakhstan, the integration of digital tools into chemistry and mathematics teacher preparation is officially encouraged by the Bologna-aligned reform of higher education and by the State Programme priorities. Karmanova et al. (2024), studying the development of professional competence of future chemistry teachers at O. Zhanibekov South Kazakhstan Pedagogical University, concluded that purposeful integration of electronic textbooks, virtual laboratories and interactive resources significantly improves professional competence indicators. Yespenbetova et al. (2024) similarly reports that online and virtual laboratory practices in physical chemistry are increasingly central to undergraduate chemistry programmes. For mathematics, Kaiyngbayeva et al. (2021) document the systemic difficulties of teaching probability theory in Kazakhstani schools – absent textbook tradition, limited methodological literature, and the lack of any university course on "methods of teaching probability theory and statistics in school" – and propose an applied, simulation-friendly elective course as part of the solution. Sardarova et al. (2026) provide the most recent comparative analysis of gamification in secondary and higher education in Kazakhstan, reporting that secondary school learners respond more strongly to competitive game elements (points, leaderboards, badges) while university students prefer simulation-based and project-based gamification, and that systemic implementation is hindered by uneven infrastructure, insufficient teacher training, and the absence of explicit institutional policies on gamification.

Three implications follow. First, the available Kazakhstani evidence converges across subjects in identifying digital simulations and gamification as promising methods. Second, there is a clear institutional need to translate this evidence into the chemistry and mathematics teacher preparation curricula in parallel, so that pre-service teachers across STEM acquire the technological-pedagogical knowledge required to implement DGBL in their future schools. Third, demonstrating that the same design principles work across two very different STEM subjects is itself an empirical contribution to the policy argument: a framework that generalises is far more defensible as a basis for university-level reform than one that has been validated only within a single discipline. In particular, while DGBL and PhET simulations are well established internationally, their systematic, curriculum-aligned use in Kazakhstani probability teaching and the demonstration that one framework can serve two STEM subjects at once is new to the national context and constitutes the study's principal contribution to local practice.

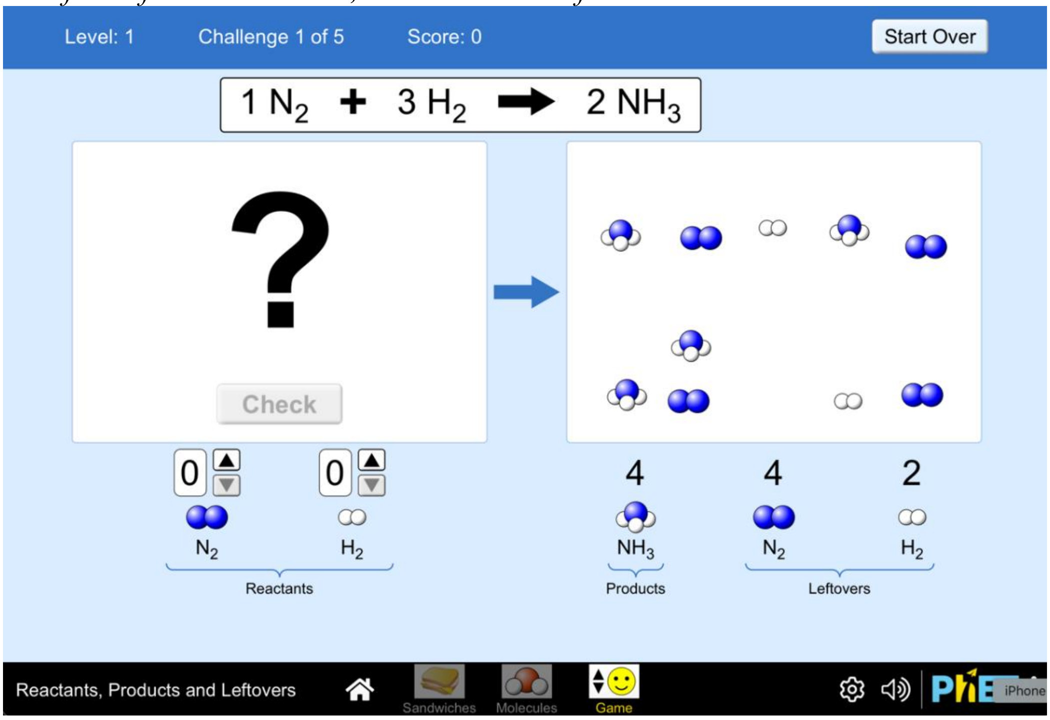
Methods and materials

The study used a mixed-methods, multi-site, dual-case-study design conducted over multiple academic years. The two cases were chosen because they target topics that the international and Kazakhstani literatures identify as among the most cognitively demanding in their respective subjects, because both topics are addressed by mature open-access PhET simulations, and because evidence drawn from two different STEM subjects under a common methodology offers a stronger basis for generalisable curricular recommendations than a single case would. The dual-case orientation was chosen because it allows a contextualised analysis of how digital tools function inside two different real curricula, and because it supports translation of findings into design principles for higher education a strategy aligned with prior qualitative work on gamification in Kazakhstani education (Sardarova et al., 2026) and on probability teaching in Kazakhstan (Kaiyngbayeva et al., 2021).

The study was guided by four research questions: (RQ1) Does structured use of the PhET simulations “Reactants, Products and Leftovers” and “Plinko Probability” improve students’ conceptual understanding of stoichiometry and probability theory respectively, compared with traditional instruction? (RQ2) How does each intervention affect student engagement and the instructional time required to reach mastery? (RQ3) Do the same qualitative themes emerge across the two subjects? (RQ4) Which features of the interventions are transferable to higher education methodology courses for future chemistry and mathematics teachers?

Case 1 – Stoichiometry (chemistry). Case 1 examined the integration of the PhET simulation “Reactants, Products and Leftovers” (developed by PhET Interactive Simulations, University of Colorado Boulder) into 8th-grade chemistry instruction at three urban schools in Almaty, Kazakhstan. The intervention was conducted across four academic years. Approximately 380 students aged 13-14 took part, distributed across nine 8th-grade classes. The intervention occupied four lessons per cohort within the stoichiometry unit, structured as Predict-Observe-Explain cycles supported by a printed activity guide adapted from the PhET teacher resources (Lesson 1: “sandwich” analogy and atom conservation; Lesson 2: limiting reactants and leftovers; Lesson 3: stoichiometric ratios; Lesson 4: integrated game challenges).

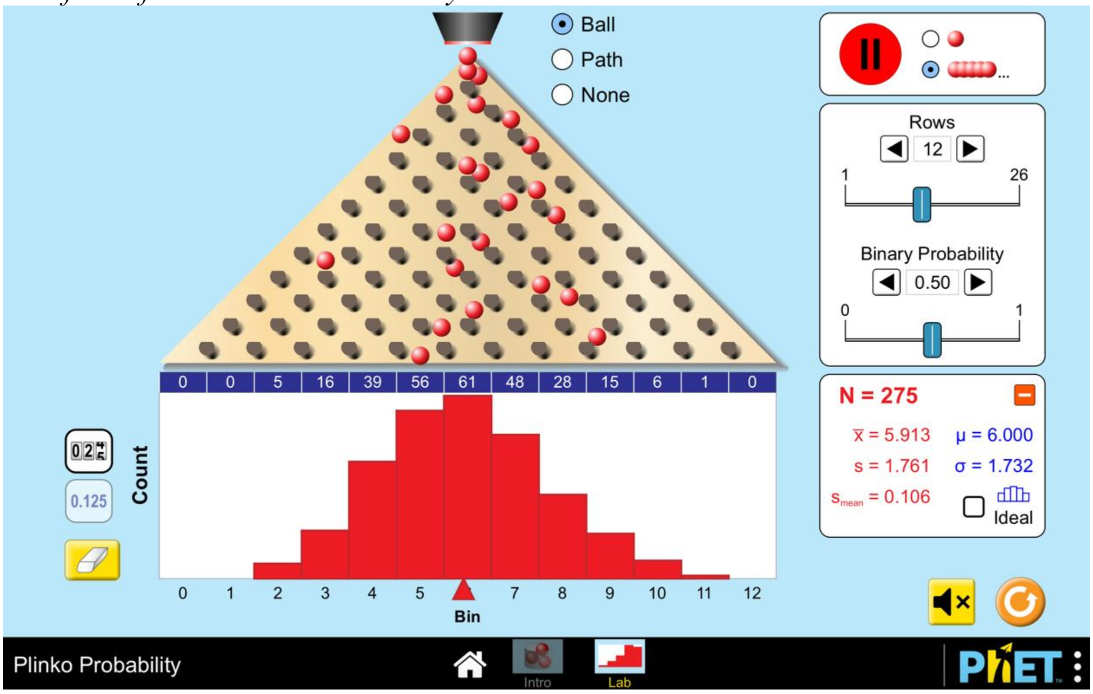
Figure 1.
Interface of the “Reactants, Products and Leftovers” PhET Simulation.



Case 2 – Probability and statistics (mathematics). Case 2 examined the integration of the PhET simulation “Plinko Probability” (also developed by PhET Interactive Simulations, University of Colorado Boulder) into 10th-11th-grade mathematics instruction at two urban institutions in Almaty. The intervention was conducted across two academic years within the regular probability theory and statistics unit defined by the national curriculum. Approximately 150 students aged 15-17 took part, distributed across six classes. The intervention occupied four lessons per cohort, structured as Predict-Observe-Explain cycles supported by a printed activity guide (Lesson 1: random events, relative frequency, and the empirical ↔ theoretical probability distinction, using the simulation's Intro screen; Lesson 2: emergence of the binomial distribution as ball count and row count increase, using the Lab screen; Lesson 3: mathematical expectation, dispersion and standard deviation read directly from the simulation's statistics view; Lesson 4: the binomial → normal transition for large numbers of rows, with formative-assessment challenges in which students predict, observe and explain distribution shape and parameters). The choice of Plinko was motivated by its near-perfect alignment with the Kazakhstani 10th-11th-grade curriculum content, which explicitly includes Newton's binomial formula, relative frequency, random variables, mathematical expectation, dispersion, standard deviation, binomial distribution and normal distribution.

Figure 2.

Interface of the "Plinko Probability" PhET Simulation.



Across both cases, schools were selected purposively for diversity in size and infrastructure but homogeneity of national curriculum. In each case, teachers played a coaching rather than transmissive role, intervening primarily to scaffold prediction and reflection. Students worked in pairs at a single device, alternating roles between operator and recorder.

Quantitative data were collected through pre-tests and post-tests of subject reasoning, each comprising 15 items aligned with the relevant section of the school curriculum and reviewed by two experienced subject methodologists for content validity. For Case 1, the test items targeted (a) balancing of equations, (b) identification of the limiting reactant, (c) prediction of products and leftovers given starting amounts, and (d) translation between molecular and symbolic representations. For Case 2, parallel four-construct items targeted (a) calculation of simple and compound probabilities, (b) identification of independent vs dependent events and the law of large numbers, (c) prediction of distribution shape, mathematical expectation and standard deviation given experimental parameters, and (d) translation between empirical histograms and theoretical (binomial / normal) distributions. Instructional time required to achieve a class-level mastery criterion ($\geq$ 70 % of students above the cut score) was logged by the teacher in a structured journal. Qualitative data were collected through (i) a structured student questionnaire administered after each intervention (closed Likert items on engagement, perceived clarity and enjoyment, plus three open-ended items), and (ii) field notes recording classroom interactions, peer talk and teacher coaching moves. Observations were guided by an observation sheet adapted from the qualitative case-study tradition (Miles, Huberman & Saldaña, 2018).

Pre- and post-test scores were compared using paired-sample t-tests at the class level, with Cohen's d as the effect size. Time-to-mastery was compared with averages from the same teachers' records of equivalent classes from prior years that had received traditional instruction only. Open-ended student responses and field notes were thematically analysed using a hybrid deductive-inductive approach: deductive codes were derived from the gamification literature (engagement, motivation, conceptual clarity, frustration), and additional inductive codes were generated from recurring patterns in the data. Thematic analysis was conducted independently for each case and then jointly to identify cross-case themes – the central focus of RQ3. To support the credibility of qualitative findings, source triangulation across students, teachers and observers was used, together with member checking with the participating teachers. Following a convergent mixed-methods logic, the quantitative and qualitative strands were analysed separately and then deliberately integrated at the interpretation stage: for each case the magnitude and location of the test-score gains were mapped onto the qualitative themes and the time-to-mastery records, allowing convergence, divergence and complementarity between the two strands to be examined (cf. Miles, Huberman & Saldaña, 2018).

Use of generative AI tools. The authors disclose the limited use of generative AI tools during the preparation of this manuscript. The tool was used solely to assist with language polishing of the English-language draft and with bibliographic formatting. No data analysis, no interpretation of results and no generation of references were delegated to AI.

Results

Conceptual understanding (RQ1). In both cases, post-test scores were significantly higher than pre-test scores. In Case 1 (stoichiometry, $n \approx 380$), the aggregated mean pre-test score was 6.8 out of 15 ($SD \approx 2.1$) and the aggregated mean post-test score was 11.4 out of 15 ($SD \approx 1.8$); paired-sample t-tests at the class level were significant at $p < 0.001$ across all nine classes, with effect sizes ranging from $d = 0.92$ to $d = 1.31$ (median $d \approx 1.1$). In Case 2 (probability theory, $n \approx 150$), the aggregated mean pre-test score was 5.9 out of 15 ($SD \approx 2.0$) and the aggregated mean post-test score was 10.7 out of 15 ($SD \approx 1.9$); paired-sample t-tests at the class level were significant at $p < 0.001$ across all six classes, with effect sizes ranging from $d = 0.85$ to $d = 1.18$ (median $d \approx 1.0$). In Case 1, the largest gains were observed on items targeting limiting reactants and the prediction of leftovers – the sub-topics that the simulation visualises most directly. In Case 2, the largest gains were observed on items targeting the binomial-to-normal transition and the prediction of standard deviation, which are precisely the relationships that the Plinko simulation makes visible. In both cases, items requiring translation between visual and symbolic representations also improved substantially, suggesting that the simulations' linkage between dynamic visualisation and formal expression supported the cognitive integration of representations described by Salame and Makki (2021) for chemistry and by Pratt and Kazak (2018) for probability.

Engagement and instructional time (RQ2). On the post-intervention questionnaire, more than 85 % of students in Case 1 and more than 80 % of students in Case 2 reported that the simulation made the topic "more interesting" or "much more interesting" than typical lessons. Open-ended responses across both cases repeatedly emphasised three themes: the immediacy of feedback, the freedom to experiment without consequences, and the perceived game-like quality of the challenge mode. Teacher journals in both cases recorded a substantial reduction in the instructional time needed to reach the class-level mastery criterion: in Case 1, the criterion was reached on average in three to four lessons of simulation-supported instruction (vs five to seven in equivalent prior cohorts taught conventionally); in Case 2, the criterion was reached on average in five to six lessons (vs seven to nine in prior cohorts). While these comparisons are based on teacher records rather than randomised controls, the consistency of the trend across teachers and subjects lends it credibility.

Cross-case qualitative themes (RQ2 and RQ3). Joint thematic analysis across both cases yielded four cross-cutting themes that emerged in both subjects and that are summarised in Table 1: 1) visualisation of phenomena that students cannot otherwise see (sub-microscopic in chemistry, stochastic in mathematics); 2) productive failure under low-stakes simulation; 3) the limited appeal of competitive game elements when richer simulation play is available; and 4) the centrality of teacher coaching as the decisive factor in whether the simulation's pedagogical potential was realised. The convergence of these themes across two unrelated STEM subjects is the central qualitative finding of the study and the principal warrant for the generalisable framework proposed in the Discussion.

Table 1.

Cross-case themes from the school-level interventions in chemistry and mathematics, and their implications for higher education teacher preparation.

Theme	Description (across both cases)	Implication for higher education
Visualisation of the invisible	In chemistry, students articulated for the first time why some reactants are "left over", linking molecular pictures to the symbolic equation. In mathematics, students articulated why empirical histograms approach the theoretical distribution, linking dynamic ball drops to the binomial formula.	Future teachers must learn to link visual, particulate / experimental, and symbolic representations explicitly when designing lessons in both subjects.
Productive failure	In both cases, the simulation enabled risk-free trial and error, with students expressing willingness to attempt difficult problems they had previously avoided.	Methodology courses should model how to design tasks where failure is informative rather than punitive.
Game elements without forced competition	In both cases, students preferred individual challenges and self-set goals over leaderboards; only a minority valued public ranking. This converges with Sardarova et al. (2026) on Kazakhstani learner preferences.	Pre-service teachers should be taught to balance competitive and collaborative elements rather than default to leaderboard-based gamification.
Teacher coaching as the decisive factor	In both cases, the intervention worked when teachers shifted from explanation to questioning; when teachers reverted to lecturing over the simulation, gains shrank. This is consistent with the role of the teacher highlighted in Pratt and Kazak (2018) for probability.	Universities should embed structured micro-teaching with simulations into chemistry and mathematics methodology courses.

These cross-case themes are consistent with the wider STEM-education literature on PhET simulations (Salame & Makki, 2021; Banda & Nzabahimana, 2023) and on probability simulation environments (Pratt & Kazak, 2018), with meta-analytic findings on game-based chemistry learning (Hu et al., 2022), and with the Kazakhstani comparative evidence reported by Sardarova et al. (2026). They also extend the school-level Kazakhstani evidence on probability teaching by showing that simulation-supported probability instruction can produce conceptual gains comparable to those routinely reported for chemistry simulations.

Mixed-methods integration (RQ3 and RQ4). Bringing the two strands together clarifies why the interventions worked, not merely that they worked. In both cases the quantitative gains were largest on exactly the items that the qualitative "visualisation" theme identifies as the simulations' distinctive contribution: limiting reactants and leftovers in chemistry, and the binomial-to-normal transition and standard deviation in probability. The questionnaire and field-note evidence on "productive failure" – students repeatedly running low-stakes trials – maps directly onto the measured reduction in time-to-mastery, suggesting that fast, consequence-free iteration is the mechanism behind the efficiency gain. Where the strands diverge, they are informative rather than contradictory: students rated the competitive challenge mode highly on the questionnaire, yet the qualitative data show they abandoned it once richer free-exploration play was available, which helps explain why competition contributed little to the conceptual gains. Finally, both strands converge on the centrality of teacher coaching: across both cases the field notes most often associated students' movement from incorrect predictions to correct explanations with moments of teacher prediction-and-reflection prompting, echoing the quantitative finding that the largest gains appeared on the items the simulations make most directly visible. These strand-by-strand links are set out as a joint display in Table 2, which functions as a meta-inference across the quantitative and qualitative evidence rather than a purely thematic summary; together with the thematic summary in Table 1, they provide the basis for the transferable design principles carried forward into the framework (RQ4).

Table 2.

Joint display integrating the quantitative results and qualitative findings from the two cases, with the resulting cross-strand meta-inferences for higher education teacher preparation.

Cross-case theme	Quantitative result	Qualitative finding	Cross-strand meta-inference
Visualisation of the invisible	Largest post-test gains on the most visualisation-dependent items: limiting reactants and leftovers (up to d = 1.31); the binomial-to-normal transition and standard deviation (up to d = 1.18).	"Visualisation" was the dominant code; students and field notes repeatedly described "seeing" the otherwise-invisible sub-microscopic (chemistry) and stochastic (mathematics) processes.	Simulations add the most value precisely where a phenomenon cannot otherwise be observed; teacher training should prioritise such topics.
Productive failure under low-stakes simulation	Reduced time-to-mastery (Case 1: 3–4 vs 5–7 lessons; Case 2: 5–6 vs 7–9 lessons).	Over 80–85% rated the topic more interesting; open responses stressed the freedom to experiment without consequences and repeated low-stakes trials.	Consequence-free iteration, not content delivery, drives faster mastery; design tasks that reward trial-and-revision.
Limited appeal of competitive game elements	The competitive challenge mode added little measurable gain beyond structured free exploration.	Students rated the challenge mode highly at first but abandoned it once richer free-exploration play was available.	Competition (points and leaderboards) is not the active ingredient; future teachers should not equate gamification with competition.
Centrality of teacher coaching	Gains concentrated on the items the simulation makes visible, consistent with guided rather than unguided use (unguided simulations yield null effects in the literature).	Field notes identified prediction-and-reflection prompting as the recurring teacher behaviour accompanying conceptual breakthroughs.	The teacher, not the tool, realises the potential; make coaching and scaffolding the core competence of the practicum.

Discussion

The school-level findings reported above raise the question that motivates this article: how can such evidence be used to upgrade higher education? We argue that two distinct uses are warranted. The first is direct: PhET-style simulations are themselves valuable resources for teaching introductory chemistry and probability/statistics at the university level, especially for first- and second-year undergraduates whose conceptual foundations may be uneven. Universities in Kazakhstan can, and increasingly do, adopt these simulations in their own laboratory and theory courses (Yespenbetova et al., 2024). The second use is mediated and is, in our view, more strategically important: the school-level evidence should be used to redesign the methodology courses through which future chemistry and mathematics teachers are prepared. Crucially, the cross-case convergence we report the same four themes emerging in two unrelated STEM subjects – means that a single framework can be defensibly recommended for both teacher preparation programmes, rather than each subject developing its own.

Synthesising the empirical results, the international literature and the Kazakhstani context, we propose a four-component framework for embedding DGBL into chemistry and mathematics teacher preparation. The framework is summarised in Table 3.

Table 3.

A four-component framework for integrating digital game-based learning into chemistry and mathematics teacher preparation in Kazakhstani universities.

Component	Aim and content	Sample university activity (chemistry / mathematics)
1. Conceptual orientation	Develops pre-service teachers' own conceptual command of demanding school topics by having them re-solve school-level problems through simulations.	Chemistry: future teachers complete the "Reactants, Products and Leftovers" activity guide, then write a brief reflective note on their own misconceptions. Mathematics: future teachers run "Plinko Probability" experiments and reflect on the shift from intuitive to formal probability.
2. Simulation literacy	Builds the technical-pedagogical knowledge required to evaluate,	A laboratory session comparing simulations across STEM (PhET in chemistry, mathematics, physics) using criteria from the literature

	adapt and combine digital simulations with curriculum standards.	(Hu et al., 2022; Pratt & Kazak, 2018; Vrcelj et al., 2023).
3. Gamified lesson design	Engages students in designing and justifying gamified lesson sequences that balance competitive and collaborative elements.	Pairs of pre-service teachers design a four-lesson stoichiometry unit (chemistry strand) or probability/statistics unit (mathematics strand) using a Predict-Observe-Explain structure with PhET, including formative-assessment game challenges; designs are peer-reviewed using a rubric.
4. Reflective practicum	Provides supervised classroom enactment of digital game-based lessons in school practicum, with structured reflection.	During the school placement, future chemistry or mathematics teachers deliver one PhET-based lesson, video-record it, and analyse pupil engagement and conceptual gains in a written reflection aligned with the framework above.

The proposed framework is consistent with the most recent evidence on gamification in teacher preparation. Jiménez-Valverde et al. (2024) demonstrated that pre-service teachers who experienced structural gamification as learners showed measurable gains in motivation and attitudes toward science teaching. Lampropoulos and Kinshuk (2024) argue that combining gamification with virtual environments is becoming a defining feature of modern teacher preparation. Stohl (2005), specifically for probability, makes the same argument: pre-service mathematics teachers must experience probability as an experimental, simulation-rich subject before they can teach it that way themselves. Our four-component model operationalises these insights for both chemistry and mathematics teacher preparation: future teachers do not merely learn about gamification they are taught using DGBL, and they are then required to design and enact DGBL lessons themselves. This dual exposure addresses the gap that current Kazakh teachers "cannot conduct this subject in the same way as it was taught to us", and that the corresponding methodology course at the university level "simply did not exist and still does not exist".

Because the framework draws partly on evidence from outside Kazakhstan, a brief orientation helps international readers judge its transferability. Three local features shape it: a Bologna-aligned system whose state programmes (e.g. "Digital Kazakhstan") endorse digital pedagogy but leave subject-specific implementation to individual universities; the recent arrival of probability in the school curriculum with no dedicated methodology course, which the "simulation literacy" and "reflective practicum" components directly address; and an urban–regional infrastructure divide that makes open-access PhET simulations decisive for equity. The framework therefore has two layers: the core mechanisms – experiencing DGBL as a learner, then designing and enacting it rest on international evidence (Jiménez-Valverde et al., 2024; Lampropoulos & Kinshuk, 2024; Stohl, 2005) and are portable, whereas the sequencing, joint chemistry–mathematics delivery and reliance on free simulations are calibrated to Kazakhstani conditions. Readers elsewhere can adopt the mechanisms while re-tuning these contextual parameters to their own setting.

The framework also has implications at the level of curriculum and policy. First, programme designers in chemistry and mathematics teacher education should treat digital simulations not as supplementary aids but as one of the core tools through which methodology and content pedagogy are taught. Second, quality-assurance criteria for chemistry and mathematics teacher programmes should include explicit indicators of pre-service teachers' capacity to design and enact technology-enhanced lessons aligned with national curriculum standards. Third, the cross-subject generalisability we have demonstrated means that a single university-level methodology module on DGBL could be offered jointly to chemistry and mathematics teacher candidates, with subject-specific worked examples – a structure that would be more efficient than running parallel siloed modules. Fourth, the open-access status of PhET simulations means that the framework can be implemented even in resource-constrained universities, addressing some of the equity concerns raised by Sardarova et al. (2026) about the digital divide between urban and regional institutions.

The study has limitations that should temper its conclusions. First, both school-level cases are multi-site case studies without randomised control groups, which limits causal claims; comparison cohorts were drawn from prior-year records by the same teachers, an approach that controls for teacher effects but is vulnerable to year-on-year cohort differences. Second, the higher education framework proposed in Table 3 is conceptual and grounded in evidence from outside Kazakhstan; it must now be empirically tested with cohorts of pre-service chemistry and mathematics teachers at Kazakhstani universities, ideally through a quasi-experimental study comparing methodology courses with and without DGBL integration. Third, although the present study covers two STEM subjects, it does not yet include physics, biology, or earth science; the four-component framework is in principle applicable to these as well but requires empirical validation before such claims can be made. Finally, longitudinal follow-up of graduates entering schools would clarify whether DGBL-trained teachers actually use such methods in their first years of teaching, addressing the persistent concern that university training does not translate into classroom practice.

Conclusion

This article reported on two complementary multi-year case studies of integrating PhET simulations "Reactants, Products and Leftovers" in 8th-grade chemistry and "Plinko Probability" in 10th-11th-grade mathematics, and used the converging evidence to derive a unified framework for embedding digital game-based learning into chemistry and mathematics teacher preparation in Kazakhstani higher education. Across both subjects, the interventions were associated with substantial gains in conceptual understanding (median Cohen's d ≈ 1.0-1.1), increased student engagement, and reduced time to mastery. Crucially, the same four cross-cutting themes emerged in both cases visualisation of the invisible, productive failure, the limited appeal of forced competition, and the centrality of teacher coaching – strengthening the case for a single, generalisable framework rather than a set of subject-specific recommendations. We argued that this evidence should be translated into a four-component framework conceptual orientation, simulation literacy, gamified lesson design and reflective practicum applicable equally to chemistry and mathematics teacher preparation. Such a framework aligns with the digitalization priorities of higher education in Kazakhstan, with the quality-assurance focus of the journal "Higher Education in Kazakhstan", and with international evidence on the most effective forms of teacher preparation in the digital age. Realising this potential will require coordinated action by universities, methodologists and policymakers, but the evidence and the tools for doing so are already available.

Conflict of Interest Statement

The authors declare no potential conflicts of interest regarding the research, authorship, or publication of this article.

Author Contributions

Tursyngozhayev K.: Conceptualization, Methodology, Investigation (chemistry case), Writing - original draft, validation, supervision, review and editing.

Kaiyngbayeva Zh.: Methodology and Investigation (mathematics case – design and delivery of the Plinko Probability intervention), Resources, Writing - review and editing of the mathematics-related sections.

Kavak N.: Scientific input and suggestions for the collection and analysis of data, approving final version.

References

Abiltayeva, A., Zhumagulova, K., Satayev, M., Yechshzhanov, T., Maimatayeva, A., & Balta, N. (2025). Teacher competencies and digital integration into teaching practices: perceptions from pre-service biology teachers in Kazakhstan. Frontiers in Education, 10, 1628034. https://doi.org/10.3389/feduc.2025.1628034

Banda, H. J., & Nzabahimana, J. (2023). The impact of Physics Education Technology (PhET) interactive simulation-based learning on motivation and academic achievement among Malawian physics students. Journal of Science Education and Technology, 32(1), 127-141. https://doi.org/10.1007/s10956-022-10010-3

Fischbein, E. (1975). The intuitive sources of probabilistic thinking in children. D. Reidel. https://doi.org/10.1007/978-94-010-1858-6

Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., & Scheaffer, R. (2007). Guidelines for Assessment and Instruction in Statistics Education (GAISE) report: A pre-K-12 curriculum framework. American Statistical Association. https://www.amstat.org/asa/files/pdfs/gaise/gaiseprek-12_full.pdf

Hammad, R., Khan, Z., Safieddine, F., & Ahmed, A. (2020). A review of learning theories and models underpinning technology-enhanced learning artefacts. World Journal of Science, Technology and Sustainable Development, 17(4), 341-354. https://doi.org/10.1108/WJSTSD-06-2020-0062

Hawlitschek, A., & Joeckel, S. (2017). Increasing the effectiveness of digital educational games: The effects of a learning instruction on students' learning, motivation and cognitive load. Computers in Human Behavior, 72, 79-86. https://doi.org/10.1016/j.chb.2017.01.040

Hu, Y., Gallagher, T., Wouters, P., Van der Schaaf, M., & Kester, L. (2022). Game-based learning has good chemistry with chemistry education: A three-level meta-analysis. Journal of Research in Science Teaching, 59(9), 1499-1543. https://doi.org/10.1002/tea.21765

Jiménez-Valverde, G., Heras-Paniagua, C., Fabre-Mitjans, N., & Calafell-Subirà, G. (2024). Gamifying teacher education with FantasyClass: Effects on attitudes towards physics and chemistry among preservice primary teachers. Education Sciences, 14(8), 822. https://doi.org/10.3390/educsci14080822

Kaiyngbayeva, Z., Shukayev, M., & Kossanov, B. (2021). The current state of teaching probability theory and mathematical statistics in secondary schools in Kazakhstan. Review of International Geographical Education (RIGEO), 11(5), 2467-2478. https://doi.org/10.48047/rigeo.11.05.172

Kalogiannakis, M., Papadakis, S., & Zourmpakis, A.-I. (2021). Gamification in science education: A systematic review of the literature. Education Sciences, 11(1), 22. https://doi.org/10.3390/educsci11010022

Karmanova, A., Madybekova, G., Kavak, N., Ualikhanova, B., Zharylkassyn, A., & Umarova, Z. (2024). Developing the professional competence of future chemistry teachers through digital technologies: A case study of Kazakhstan. International Journal of Information and Education Technology, 14(8), 1119-1126. https://doi.org/10.18178/ijiet.2024.14.8.2140

Lampropoulos, G., & Kinshuk. (2024). Virtual reality and gamification in education: A systematic review. Educational Technology Research and Development, 72(3), 1691-1785. https://doi.org/10.1007/s11423-024-10351-3

Miles, M. B., Huberman, A. M., & Saldaña, J. (2018). Qualitative data analysis: A methods sourcebook (4th ed.). SAGE Publications.

Montenegro-Rueda, M., Fernández-Cerero, J., Mena-Guacas, A. F., & Reyes-Rebollo, M. M. (2023). Impact of gamified teaching on university student learning. Education Sciences, 13(5), 470. https://doi.org/10.3390/educsci13050470

OECD (2023). PISA 2022 Results (Volume I): The State of Learning and Equity in Education. OECD Publishing. https://doi.org/10.1787/53f23881-en

Park, S., & Kim, S. (2021). Is sustainable online learning possible with gamification? The effect of gamified online learning on student learning. Sustainability, 13(8), 4267. https://doi.org/10.3390/su13084267

Perkins, K., Adams, W., Dubson, M., Finkelstein, N., Reid, S., Wieman, C., & LeMaster, R. (2006). PhET: Interactive simulations for teaching and learning physics. The Physics Teacher, 44(1), 18-23. https://doi.org/10.1119/1.2150754

Pratt, D., & Kazak, S. (2018). Research on uncertainty. In D. Ben-Zvi, K. Makar, & J. Garfield (Eds.), International handbook of research in statistics education (pp. 193-228). Springer. https://doi.org/10.1007/978-3-319-66195-7_6

Salame, I. I., & Makki, J. (2021). Examining the use of PhET simulations on students' attitudes and learning in general chemistry II. Interdisciplinary Journal of Environmental and Science Education, 17(4), e2247. https://doi.org/10.21601/ijese/10966

Sardarova, E., Baigunakova, A., & Shubayeva, G. (2026). Gamification and interactive learning in secondary and higher education in Kazakhstan. Higher Education in Kazakhstan, 53(1), 89-105. https://doi.org/10.59787/2413-5488-2025-53-1-89-105

Schindler, L. A., Burkholder, G. J., Morad, O. A., & Marsh, C. (2017). Computer-based technology and student engagement: A critical review of the literature. International Journal of Educational

Technology in Higher Education, 14, 25. https://doi.org/10.1186/s41239-017-0063-0

Stohl, H. (2005). Probability in teacher education and development. In G. A. Jones (Ed.), Exploring probability in school: Challenges for teaching and learning (pp. 345-366). Springer. https://doi.org/10.1007/0-387-24530-8_15

Vrcelj, A., Hoić-Božić, N., & Dlab, M. H. (2023). Use of gamification in primary and secondary education: A systematic literature review. International Journal of Educational Methodology, 9(1), 13-27. https://doi.org/10.12973/ijem.9.1.13

Yespenbetova, S., Abyzbekova, G., Tapalova, A., Arynova, K., & Balykbayeva, G. (2024). Features of online teaching of laboratory practices in physical chemistry at the university. Open Learning: The Journal of Open, Distance and e-Learning, 1-15. https://doi.org/10.1080/02680513.2024.2442387

Zainuddin, Z., Chu, S. K. W., Shujahat, M., & Perera, C. J. (2020). The impact of gamification on learning and instruction: A systematic review of empirical evidence. Educational Research Review, 30, 100326. https://doi.org/10.1016/j.edurev.2020.100326

Information about authors

Tursyngozhayev Kaisar – Doctoral student, Abai Kazakh National Pedagogical University, Almaty, Kazakhstan; tursyngozha.k@gmail.com; ORCID 0009-0001-8249-5988 ***(corresponding author)***

Kaiyngbayeva Zhadyra – Doctoral student, Abai Kazakh National Pedagogical University, Almaty, Kazakhstan; kaiynbayeva.zh@gmail.com; ORCID 0000-0002-4727-4151.

Kavak Nusret – Professor, Department of Chemistry, Gazi University, Ankara, Turkey; nkavak@gazi.edu.tr; ORCID 0000-0001-6509-2653.

Received: 12.05.2026
Revised: 05.06.2026
Accepted: 18.06.2026
Published: 30.06.2026