

INTEGRATING MATHEMATICAL ACTIVITIES THROUGH GAMES TO ESTABLISH A MATHEMATICAL FOUNDATION AND PROMOTE CRITICAL THINKING IN PRE-ELEMENTARY CHILDREN

Nguyen Duy Hai, Nguyen Thi Quynh Hoa, Tran Tuan Long*, Khong Do Thanh Huyen
and Chu Dinh Minh

*School of Mathematics and Information Technology, Hanoi National University of Education,
Hanoi, Vietnam*

*Corresponding author: Tran Tuan Long, e-mail: longtt@longit.id.vn

Received: January 17, 2026. Revised: March 14, 2024. Accepted: April 06, 2026.

Abstract. Developing mathematical skills and basic critical thinking for young children is an important goal in preschool education, yet current studies often underestimate the connection between these two factors. This study proposes a web-based gamification method that combines mathematical skills training with the detection of critical thinking expressions through interactive online learning games. To evaluate the system's effectiveness, a two-week controlled experiment was conducted with 39 preschool children aged 4 to 5 years old. Participants were divided into an experimental group (n=16) using the digital platform and a control group (n=23) using traditional paper-based methods. The application, developed using HTML, CSS, JavaScript, and PHP, tracks student performance indicators and maps them against the FRISCO critical thinking framework. Independent t-test results indicated a statistically significant improvement in mathematical scores for the experimental group compared to the control group. Furthermore, the system's automated assessment of critical thinking showed a strong correlation with expert evaluations in the "Focus" ($r=0.970$) and "Clear" ($r=0.984$) criteria. Ultimately, this research presents a practical technological architecture that supports teachers in continuously monitoring early mathematical progress and creating a highly engaging learning environment.

Keywords: Mathematical skills, preschool children, critical thinking, FRISCO, gamification

1. Introduction

One of the valuable skills that the education system can develop in students is critical thinking (O'Reilly, 2022). However, critical thinking is not innate; it is formed and developed over time and through practice (Kuhn, 2016; Peter, 2012; Snyder & Snyder, 2008). Critical thinking plays an important role from an early age because as children acquire new knowledge, they also need to be able to identify and avoid misinformation (Brousseau-Liard, 2017). Studies by Heyman and Strasser have shown that critical thinking can be developed in children as young as three years old (Heyman, 2008; Strasser & Bresson, 2017). In the context of modern global education, critical thinking and mathematical skills are considered essential foundations for helping students develop logical thinking, creativity, and effective problem-solving abilities. When students first get acquainted with mathematics, mastering basic knowledge and practicing critical thinking are decisive factors for their confidence and ability to learn later (Sachdeva & Eggen, 2021). Mathematics education for young children includes providing stimulating activities and learning environments that provide young children with experiences that expand their knowledge and develop their mathematical concepts and skills (Björklund et al., 2020). The

question of how, and to what extent, pre-primary school children should be introduced to mathematics has been widely discussed in early childhood mathematics education, including in the POEM conference series (Carlsen et al., 2018). With the continuous development of technology, a series of IT application solutions in teaching open up a new perspective on mathematics for young children such as symmetry transformation (Fletcher & Ginsburg, 2016; Sinclair, 2018). In the preschool stage, play is a dynamic learning environment (Sinclair, 2018), where children's learning often takes place (Reikerås, 2020; Uğurel & Moralı, 2008). Regardless of the stage, games exist almost in every aspect of life; they originate from life, develop, and transform according to needs. Therefore, game-based learning creates an environment for exploration and development that is suitable for children's ability to absorb knowledge and attention spans (Aksoy, 2014; Güneş, 2015).

Therefore, this study proposes a method to support the assessment of critical thinking ability and practice of mathematical skills based on the application of web-based gamification. This solution not only helps to collect the necessary indicators flexibly and naturally, but also creates a friendly learning environment, reducing the pressure on children when they are new to mathematics. The remainder of this paper is organized as follows. Section 2 reviews studies on game-based learning and technology-supported mathematics education. Section 3 presents the proposed method and system design. Section 4 describes the experiment and discusses the results. Section 5 concludes the study and suggests future research directions.

2. Literature review

Studies on games in learning have demonstrated that this is a suitable method to support preschool children in learning mathematics. Mathematical games can effectively attract children's participation and support the development of their mathematical abilities and skills (Cohrssen & Niklas, 2019). In teaching practice, gamification strategies are implemented by educators through a variety of technological solutions, aiming to integrate game elements into mathematics learning activities.

Teachers often choose applications (apps) designed to support students in practicing basic mathematical skills, especially because these tools can provide interactive, engaging, and repeated practice opportunities for young learners (Mera et al., 2019). Integrating tools with gamification elements into the teaching process helps children enhance mathematical skills through interactive activities. However, the study also points out a major limitation: these applications often lack feedback mechanisms when children answer incorrectly. Sukstrienwong presented a teaching method through the use of the role-playing game *Animo Math*, which aims to support children in learning mathematics in an interesting and effective way while developing mathematical skills (Sukstrienwong, 2018). However, focusing too much on the cartoon characters in the game can distract or overwhelm children, reducing their focus on the learning goal and the effectiveness of mathematical skill improvement.

Teachers' approaches to early mathematics instruction often combine traditional classroom activities with game-based learning to help children develop basic mathematical skills, such as number recognition, while enhancing interaction and engagement in the learning process (Orr et al., 2015). However, each child has a different learning speed. Although the curriculum provides materials, it still cannot meet each child's learning level, especially in an environment with great diversity in learning needs. A gamified mathematics teaching model using EDI software for preschool children can support the development of basic mathematical skills and cognitive abilities through an interactive and engaging learning environment (Martins & Silva, 2017). In addition, a limitation of the study is the lack of sound, which hindered children's interaction with the software.

Teaching strategies implemented through online learning platforms such as Kahoot, Educaplay, and Genially, which integrate gamification elements, can positively support the development of logical thinking and mathematical skills in preschool children (Mejia & Sargent, 2023). A notable limitation of the study is that specific gamification elements - such as the reward system, challenge level, or feedback mechanism - have not been clearly identified. The lack of in-depth analysis of each component of the game makes it difficult to evaluate the effectiveness of each element, thereby affecting the ability to improve, adjust, or replicate the gamification application model in preschool education environments. Technology-integrated teaching models can support teachers in designing learning activities that promote students' critical thinking skills (Mejia & Sargent, 2023). However, assessing students' critical thinking ability through this technology can be difficult due to the lack of a suitable assessment tool with clear standards.

3. Proposed method

The study proposes a method to integrate mathematical activities through games to establish a mathematical foundation. Consequently, it assesses students' mathematical skills in detail and provides indicators of critical thinking tendencies through web-based learning games. The proposed method of the study is illustrated in Figure 1.



Figure 1. General diagram

Survey collection: The survey was conducted in March 2024 across five public kindergartens in Hanoi, Vietnam. Participants were 35 preschool teachers currently teaching children aged 4 to 5, selected through purposive sampling based on two criteria: at least two years of classroom experience with this age group, and current responsibility for mathematics instruction. The survey instrument consisted of 24 closed-ended items using a five-point Likert scale (1 = strongly disagree, 5 = strongly agree), organized into four thematic subscales: (1) mathematical content knowledge to be conveyed (7 items); (2) methods of assessing children's learning outcomes (6 items); (3) observable expressions of critical thinking in the classroom (6 items); and (4) challenges encountered when assessing critical thinking using traditional approaches (5 items). The instrument was piloted with six teachers outside the main sample before full administration, and items were refined accordingly. Internal consistency was verified using Cronbach's alpha: coefficients for the four subscales were 0.82, 0.79, 0.81, and 0.77 respectively, all exceeding the accepted threshold of 0.70 and confirming satisfactory reliability. The survey was distributed in printed form during a scheduled professional development session and collected on the same day, achieving a 100% response rate.

Survey analysis: Descriptive statistics were computed for all items. Regarding mathematical content and technology preferences, 80% of teachers ($n = 28/35$) rated touch-based interactions such as touching, swiping, and dragging as highly effective in attracting children's attention ($M = 4.18$, $SD = 0.61$), and 86% ($n = 30/35$) agreed that interactive exercises promote more active learning ($M = 4.32$, $SD = 0.55$). In terms of interface design, 88% of teachers emphasized the importance of immediate corrective feedback ($M = 4.41$, $SD = 0.52$), followed by clear button design (70%, $M = 3.97$, $SD = 0.68$) and leveled tasks (76%, $M = 4.10$, $SD = 0.59$). The most commonly taught content areas for ages 4-5 included geometric shape recognition, size comparison, counting and number recognition, and positional concepts. Regarding observable expressions of critical thinking in the classroom, teachers most frequently reported children asking "why" questions during tasks ($M = 4.05$, $SD = 0.63$), self-correcting errors before seeking adult help ($M = 3.89$, $SD = 0.71$), and selecting between alternative solution strategies ($M = 3.74$, $SD = 0.78$). Regarding challenges of assessing critical thinking through traditional

methods, 91% of teachers ($n = 32/35$) agreed it is difficult to simultaneously observe and document critical thinking behaviors while managing the classroom ($M = 4.37$, $SD = 0.58$), and 83% ($n = 29/35$) reported that paper-based tools are insufficient for capturing the process-level behaviors that indicate emerging critical thinking ($M = 4.14$, $SD = 0.64$). These findings collectively provide the empirical basis for the web-based gamification approach proposed in this study.

Design: Based on the survey results, the study builds a rich system of learning tasks associated with mathematical content for preschool children such as counting numbers, recognizing shapes, comparing sizes and determining positions. The tasks were designed flexibly in terms of format and difficulty level, and they integrated interactive elements and immediate feedback to enhance children's learning experiences. These requirements are implemented through games such as pouring water, bingo, and jigsaw puzzles, with variable content and gameplay, suitable for many groups of children. The design model meets practical needs and supports teachers in organizing and adjusting effective teaching activities.

Control mechanism: Games can be fun for children, but they also have the potential to be distracting. Therefore, this study proposes a web-based gamification method to support the practice of mathematical skills and the assessment of early behavioral indicators associated with critical thinking tendencies. Game participation is always supervised by teachers or parents, and students are only allowed to participate in assigned activities, rather than initiating new games themselves.

Interactive learning: The initial stages of learning mathematics are often difficult due to the abstract nature of knowledge and young children's short attention span, which can easily lead to boredom and reduced learning efficiency. The study proposes to develop short, interactive and flexible tasks, combining three game modes: independent, collaborative and competitive, to create an attractive, proactive learning environment and increase children's motivation to learn.

Topic: The selected math topics focus on familiar content, suitable for preschool children's perception and teachers' teaching needs. Priority is given to topics such as recognizing and distinguishing numbers, classifying by color and shape. These are core components in preschool programs, which are easy to integrate into interactive activities and highly feasible when turned into games, helping children access easily and learn more actively.

Reward factor: To kick-start the learning journey with excitement, gamification tools need to capture children's attention right from the start through strong visual elements such as contrasting colors, vivid images, interactive effects, and engaging sounds. This combination stimulates the senses, arouses curiosity, and motivates children to participate in learning. However, the initial attraction is only the beginning. To maintain focus, the method continues to apply game elements such as an incentive reward system, diverse tasks, innovative content, and a reasonable level of difficulty, helping children always feel at ease, motivated, and develop sustainably.

Detailed review: Based on the teachers' learning plans and assigned tasks to students, the method develops a detailed assessment tool for the process of students performing tasks to provide information on scores and students' performance process. Instead of focusing only on the final result of "right" or "wrong", the tool analyzes each step of students' performance throughout the task, the time spent on the task, the number of attempts, the number of corrections, etc. Thus, it provides instant information for teachers to be able to change the task to suit the developmental level of students.

System architecture: The proposed system is built on a three-tier web architecture designed to be lightweight, accessible, and deployable in standard preschool classroom environments without requiring specialized hardware. The Frontend (User Interface) is developed using HTML5, CSS3, and JavaScript, employing high-contrast colors and large interactive touch targets to sustain the attention of children aged 4–5 while replacing text-heavy instructions with intuitive visual and auditory cues. The Backend (Logic and Database Layer) is managed via PHP and a MySQL database, which records granular behavioral data, including interaction sequences, error pathways, and response latencies, to serve as inputs for the heuristic FRISCO assessment. A dedicated Teacher Dashboard visualizes these process-

level indicators, such as self-correction frequency and completion trajectories, enabling educators to identify individual learning needs and adjust task difficulty in a data-informed manner. Given the young age of the learners, the system employs a rule-based heuristic approach to map student behaviors onto FRISCO indicators (Ennis, 2011). Instead of using advanced machine learning models or complex sensing hardware. Specific behavioral mapping rules for each criterion are detailed in

Table 1 below.

Table 1. FRISCO critical thinking indicators

FRISCO Criterion	Original Definition	Web-System Behavioral Cue (Rule-Based Tracking)
Focus (F)	Understand information and distinguish relevant/irrelevant data.	Action: Time spent actively engaging with the task interface vs. idle time. Metric: Rapid identification of the correct target shape/number without clicking irrelevant background elements.
Reason (R)	Give reasons based on evidence for each step.	Action: Sequential logic in puzzle-solving. Metric: Successfully dragging and dropping items in the correct order (e.g., smallest to largest) without erratic, random placements
Inference (I)	Draw correct conclusions and provide appropriate reasons.	Action: Applying a learned rule to a new level. Metric: High accuracy rate on the first attempt of a new, slightly harder task level, indicating the child inferred the underlying mathematical rule.
Clear (C)	Explain terms and give examples of similar cases.	Action: Adapted for preschool level — recognizing visual patterns. Metric: Successfully grouping similar items (e.g., all blue triangles), demonstrating a non-verbal understanding of the classification concept.
Situation (S)	Understand and apply knowledge within the context of the problem.	Action: Understanding the context of the game scenario. Metric: Correctly choosing the right tool or operation for the given task context (e.g., selecting the "pouring" action in a volume comparison game).
Overview (O)	Study or review the process thoroughly from beginning to end.	Action: Self-correction and review behavior. Metric: The number of times a student corrects a mistake before submitting the final answer, combined with overall task completion time relative to the session average.

Firstly, regarding the theoretical foundation and behavioral concretization: The study deconstructed core cognitive skills from the Delphi project consensus report by Facione (Facione, 1990) to serve as the basis for establishing indicators. Due to the specific psychological limitations of preschool children (4-5 years old) in their ability to express complex language, abstract thinking concepts were transformed into corresponding behavioral proxies in the digital environment. Specifically, the Analytical skill was concretized into the Focus criterion (F) through the ability to filter noise and focus on the goal; the Interpretation skill was mapped to the Clear criterion (C) through the behavior of recognizing visual patterns; In particular, self-correction is the empirical manifestation of Self-Regulation skills—a high-level component confirmed by 46 international experts in the Facione report (Facione, 1990). Secondly, regarding quantitative content assessment: The finalized framework was independently assessed by two early childhood education experts with over 5 years of practical experience, following the standard procedure of Polit and Beck (Polit & Beck, 2006). The experts evaluated the developmental suitability of each indicator on the 4-point Likert scale (from 1: unsuitable to 4: very suitable). The application of the 4-point scale is an important methodological technique to eliminate neutral or hesitant responses, forcing experts to make a definitive assessment of the compatibility of digital behavior with critical thinking criteria. The evaluation results show that the core indicators achieved absolute consensus, contributing to an overall score (S-CVI/Ave) of 0.92, confirming the content validity of the toolkit at an excellent level according to international standards.

Thirdly, regarding empirical validation and the limitations of substitutes: The reliability of the validation process is strengthened by a Kappa coefficient of 0.76, demonstrating substantive agreement between the system's recorded data and the experts' direct observations. However, for higher-order criteria such as Reason (R) and Situation (S) - which require linguistic expression in the original model - the current system can only record data through behavioral interaction patterns. These substitutes inherently have certain limitations for children aged 4-5, explaining the lower correlation coefficients for these criteria in the experimental results. Conversely, criteria such as Focus (F) and Clear (C), thanks to their ability to directly observe nonverbal behavioral data, showed a stronger correlation between the system's assessment and expert opinion.

Pedagogical Scenarios: The system supports three teacher-supervised interaction modes-individual, collaborative, and competitive play-each designed to serve distinct instructional purposes and assessment workflows.

Scenario 1: Independent Exploration and Competency Profiling (Individual Play) involves personalized tasks like "water pouring," where the system establishes a visual convention mapping numerical units to water levels, effectively making abstract numbers tangible as physical quantities. This allows children to internalize mathematical operations through trial and error at their own cognitive pace, while providing teachers with behavioral indicators - such as interaction response times and total click counts-to identify hesitation or mastery as a scientific basis for targeted intervention.

Scenario 2: Peer Learning and Inferential Reasoning (Collaborative Play) pairs students to solve logic challenges within numerical matrices, which can be configured for varying difficulty - from finding single pairs to identifying combinations of 1, 2, ..., n numbers whose sum equals a target value M. This challenge requires children to engage in active discussion and collaboratively construct strategies, allowing the system to track data for the Inference (I) indicator while the teacher assesses teamwork and communication skills.

Scenario 3: Formative Assessment through Competitive Play, exemplified by mathematical Bingo, requires rapid knowledge retrieval under time pressure to surface the Focus (F) and Clear (C) behavioral indicators within the FRISCO framework. The system automatically provides real-time performance reports to the teacher's dashboard, serving as a scientific basis for adjusting subsequent teaching plans to align closely with actual whole-class comprehension levels.

4. Experiment

4.1. Target

The experiment was conducted to achieve two main objectives: first, to evaluate the effectiveness of a web-based digital learning method in improving children's mathematical skills through gamification; and second, to evaluate the method's ability to detect and record behavioral indicators associated with critical thinking tendencies during the learning process, thereby contributing to supporting objective and continuous assessment of skills. It should be noted that, given the age group of 4-5-year-old children, this study does not aim to fully assess all FRISCO dimensions at the same cognitive level as older learners. Rather, it focuses on identifying early behavioral manifestations that may serve as observable proxies for emerging critical thinking tendencies, as supported by prior research indicating that early forms of critical thinking can emerge in children as young as three years old (Heyman, 2008; Strasser & Bresson, 2017).

4.2. Experimental procedure

To perform the two-objective evaluation of the proposed method, the necessary work is deployed including four steps illustrated in

Figure 2.

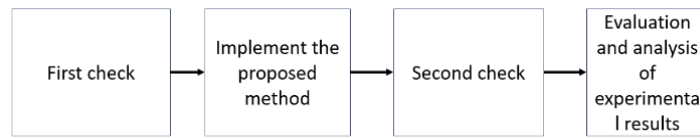


Figure 2. Experimental procedure

Test 1: To ensure similarity in initial mathematical ability, students took a paper-based entrance test. Based on the test results, students were divided into two groups of similar levels: Group 1 (experimental) with 16 students) and Group 2 (control) with 23 students. In the experimental group, the distribution of performance levels was: 31.25% (5 students) achieved good or excellent performance; 43.75% (7 students) achieved average performance; and 25% (4 students) achieved poor performance. In the control group, the corresponding rates were: 30.43% (7 students) achieved good or excellent performance; 43.48% (10 students) achieved average performance; and 26.09% (6 students) achieved poor performance. The grouping was done carefully to ensure a balance of skill levels between the two groups before conducting the experiment, in order to compare the effectiveness of the proposed method (experimental group) and the traditional method (control group) objectively. We acknowledge that the sample size of 39 children, with 16 in the experimental group and 23 in the control group, is relatively small. This is a recognized limitation of the current study, arising from constraints related to the number of available participants at the cooperating preschool institution during the study period. The authors are aware that a larger and more balanced sample would strengthen the statistical power and generalizability of the findings. Accordingly, the results reported in this study should be interpreted as preliminary and indicative rather than definitive, and future research will aim to replicate and expand the experiment with a larger sample.

Implementation of the proposed method: After the initial assessment and grouping phase, the study conducted an intervention phase lasting two weeks, with a frequency of 3 sessions per week (totaling 6 sessions). During this phase, the experimental group used the proposed web-based gamification system under teacher supervision, while the control group received conventional instruction using traditional teaching methods without digital game-based tools. All sessions for both groups were conducted in the same preschool environment to minimize confounding variables. We acknowledge that the two-week intervention period is relatively short for evaluating changes in complex cognitive abilities such as critical thinking. This duration was constrained by the schedule of the cooperating institution and the practical limitations of conducting extended studies with very young children. Therefore, the findings related to mathematical skill improvement should be regarded as short-term effects, and the behavioral indicators identified by the system represent early manifestations rather than fully developed critical thinking competencies. Extended longitudinal studies are recommended for future research to assess the lasting effects.

Second Test: After the two-week implementation phase, an exit test (second test) was conducted synchronously for both groups. The exit test was designed with the same difficulty level and knowledge content coverage as the entry test, ensuring comparability of results between the two testing time points and between the two groups.

Evaluation and analysis of experimental results: Performance data from the exit test of all 39 students were collected. For the experimental group, the test data were automatically evaluated by the system. For the control group, the written test was manually graded by teachers according to unified grading guidelines. Regarding the critical thinking assessment component, the system's output was compared against independent evaluations conducted by two mathematics education specialists, each with more than five years of professional experience in early childhood mathematics education. The two experts assessed the same student task records independently using the FRISCO framework indicators,

focusing on the behavioral dimensions observable within game-based interactions (specifically, Focus and Clear criteria, which are more amenable to behavioral observation in this age group). To ensure the reliability of expert judgments, inter-rater agreement was calculated. The Cohen's Kappa coefficient between the two experts was $\kappa = 0.76$, indicating substantial agreement and confirming the reliability of the expert assessment as a reference standard for comparison with the system output. We further acknowledge that for certain FRISCO criteria - particularly Reason (R) and Situation (S)-the automatic detection is inherently limited in the context of game-based interaction with young children, as these criteria require verbal explanation and metacognitive awareness that cannot be fully captured through behavioral interaction data alone. The correlation results for these criteria are accordingly lower and should be interpreted with caution.

4.3. Experimental results

Math skills assessment results: To evaluate the effectiveness in improving students' math skills, the study conducted t-tests as follows:

Evaluating the difference in mean scores between the two test attempts of each paired group: A paired t-test was performed to assess the difference in mean scores between the first test attempt (Time 1) and the second test attempt (Time 2) in each group.

Table 2. Paired sample test results of group 1 (experimental group)

		Paired Differences						
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference (Lower)	95% Confidence Interval of the Difference (Upper)	t	df
Pair 1	Pair 2	-3.89375	1.97247	.49312	-4.94480	-2.84270	-7.896	15

Table 3. Paired sample test results of group 2 (control group)

		Paired Differences						
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference (Lower)	95% Confidence Interval of the Difference (Upper)	t	df
Pair 1	Pair 2	-1.73696	2.00081	.41720	-2.60217	-.87174	-4.163	22

The results from the paired t-test for both groups showed a significant improvement in scores from time 1 to time 2. This shows that both the intervention method in the experimental group and the conventional method in the control group contributed to the improvement of students' mathematics skills after a period of study.

Assessing the difference in mean scores between the two experimental groups: To compare the relative effectiveness between the two methods, an independent t-test was conducted to determine the difference in mean scores between the experimental group (Group 1) and the control group (Group 2) in the second test (Time 2)

Table 4. Independent sample t-test results

	Levene's Test for Equality of Variances	t-test for Equality of Means							
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% CI (Lower)
Equal variances assumed	.742	.394	2.618	37	.013	1.53859	.58768	.34784	2.72934

Equal variances not assumed		2.525	28.127	.017	1.53859	.60928	.29079	2.78639
-----------------------------	--	-------	--------	------	---------	--------	--------	---------

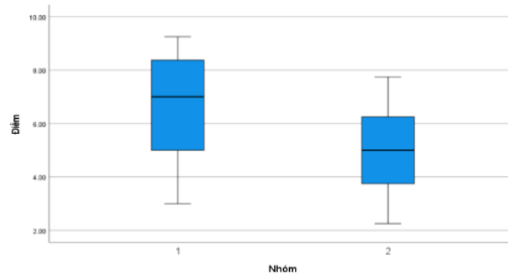


Figure 3. The average scores of the two groups

The results of the independent t-test showed that there was a statistically significant difference in the mean scores between the experimental group and the control group in the second test. The experimental group (Group 1) achieved a higher mean score than the control group (Group 2). This suggests that the proposed method (applied to the experimental group) is more effective than the conventional method (applied to the control group) in improving mathematical skills. The mean score difference between the two groups underscores the effectiveness of the proposed method.

Evaluation results of the proposed method's ability to detect students' critical thinking ability level: To evaluate the accuracy of the proposed method in identifying mathematical ability indicators according to the FRISCO assessment framework, the research team compared the ability assessment results provided by the system with the professional assessment of experts in the field of mathematics education. Specifically, students' assignments were analyzed by the system and labeled with corresponding abilities. At the same time, the same assignments were also independently assessed by experts according to the same ability framework. The comparison results between the two assessment sources are summarized and presented in Figure 4.

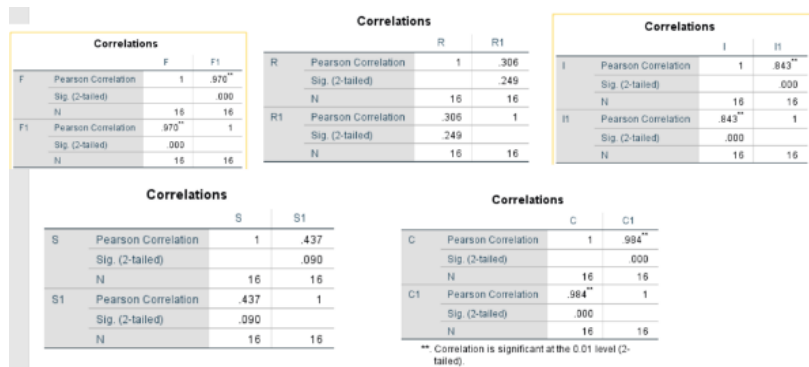


Figure 4. Comparison of critical thinking ability assessment results between the proposed method and experts

The analysis results show that the proposed method demonstrates strong alignment with expert assessments for certain criteria. In particular, for Focus (F) and Clear (C), the correlation coefficients reached 0.970 and 0.984 respectively, indicating a very strong relationship between the system's assessments and expert evaluations. These two criteria are more directly observable through behavioral interaction data (such as attention consistency and task completion patterns), which explains the high alignment. However, criteria such as Reason (R) and Situation (S) yielded lower correlation coefficients of 0.437 and 0.306 respectively, reflecting only moderate agreement. This is an expected limitation: these criteria require children to articulate reasoning and respond to contextual nuances - competencies that

are difficult to capture through game interaction data alone, particularly in this age group. The absence of the Overview (O) criterion from the correlation figure is acknowledged; this criterion was assessed qualitatively but could not be reliably operationalized into system-detectable behavioral indicators within the current implementation and is a target for future development. Although the current system has limitations in detecting higher-order critical thinking indicators, the initial results demonstrate meaningful potential for the automated behavioral assessment of early critical thinking manifestations in young children. The system will continue to be researched, refined, and expanded in subsequent stages to improve detection accuracy, address the identified criterion gaps, and better serve pedagogical goals in ability-oriented early childhood education.

5. Conclusions

The study has developed a web-based, game-based math learning method for children preparing to enter first grade. Math calculations are presented in vivid, easy-to-understand images, helping children see and clearly imagine the relationship between numbers. Consequently, children can easily absorb knowledge and develop mathematical thinking. The game system has three different interaction modes: individual play, collaborative play, and competitive play. In individual mode, children can explore and practice at their own pace. In collaborative mode, they learn to share, cooperate and help each other to solve problems. In competitive mode, children will be able to try their hand at interesting challenges, helping to increase their interest and competitive spirit when learning math. Initial results show that the game system not only helps children learn better but also makes learning math more fun. In addition, the game also helps children practice many other important skills such as logical thinking, teamwork and self-study habits. This is a new approach, helping to bring technology into teaching math to young children in an effective and creative way.

Note for contributor: *Nguyen Duy Hai is a teaching assistant, Nguyen Thi Quynh Hoa is a lecturer, and Tran Tuan Long, Khong Do Thanh Huyen, and Chu Dinh Minh are students at the School of Mathematics and Information Technology, Hanoi National University of Education, Vietnam. Author 1: conceptualization, methodology, supervision; author 2: writing original draft, review & editing; author 3: software, system design; author 4: data analysis, visualization; author 5: data collection, writing.*

Conflict of interest: *The authors declare no conflict of interest.*

Acknowledgments: *The authors thank the teachers who participated in the survey and the preschool in Hanoi for their support during the experiment.*

REFERENCES

- Aksoy, N. (2014). *Effects of digital game-based mathematics teaching on 6th grades students' achievement, motivation, attitude and self-efficacy*. Gazi University.
- Björklund, C., van den Heuvel-Panhuizen, M., & Kullberg, A. (2020). Research on early childhood mathematics teaching and learning. *ZDM*, 52(4), 607–619. <https://doi.org/10.1007/s11858-020-01177-3>
- Brousseau-Liard, P. (2017). The roots of critical thinking: Selective learning strategies in childhood and their implications. *Canadian Psychology/psychologie canadienne*, 58, 263.
- Carlsen, M., Erfjord, I., & Hundeland, P. (2018). Mathematics education in the early years. *Results from the POEM4 Conference*. Springer.
- Cohrssen, C., & Niklas, F. (2019). Using mathematics games in preschool settings to support the development of children's numeracy skills. *International Journal of Early Years Education*, 27(3), 322-339.
- Ennis, R. (2011). The nature of critical thinking: An outline of critical thinking dispositions and abilities. *University of Illinois*, 2, 1-8.

- Facione, P. (1990). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction (The Delphi Report)*.
- Fletcher, N., & Ginsburg, H. (2016). Teaching symmetry in the early childhood classroom: Using software to facilitate discourse and promote understanding of symmetric transformations. *13th International Congress on Mathematical Education*.
- Güneş, F. (2015). Game-based learning approach. *Turkish Studies*, 10, 773-786.
- Heyman, G. (2008). Children's Critical Thinking When Learning From Others. *Current Directions in Psychological Science*, 17(5), 344-347. <https://doi.org/10.1111/j.1467-8721.2008.00603.x>
- Kuhn, D. (2016). A Role for Reasoning in a Dialogic Approach to Critical Thinking. *Topoi*, 37, 121-128. <https://doi.org/10.1007/s11245-016-9373-4>
- Martins, E., & Silva, O. (2017). Mathematics for early childhood education: A model of mathematics teaching through EDI software. *IOSR Journal of Research & Method in Education*, 07(1), 64-78. <https://doi.org/10.9790/7388-0701056478>
- Mejia, M., & Sargent, J. (2023). Leveraging technology to develop students' critical thinking skills. *Journal of Educational Technology Systems*, 51(4), 393-418.
- Mera, C., Ruiz, G., Román, B., Aragón, E., & Navarro, J. (2019). Apps para el aprendizaje de las matemáticas en educación infantil [Apps for learning mathematics in early childhood education]. *Revista INFAD de Psicología. International Journal of Developmental and Educational Psychology*, 3, 121-132.
- O'Reilly, C., Devitt, A., & Hayes, N. (2022). Critical thinking in the preschool classroom - A systematic literature review. *Thinking Skills and Creativity*, 46, 101110. <https://doi.org/10.1016/j.tsc.2022.101110>
- Orr, J., Flannery, L., Presser, A., Vahey, P., & Latimore, S. (2015). Early Math with Gracie & Friends™ Demo: App-infused curriculum and teacher support for preschool. *Proceedings of the 14th International Conference on Interaction Design and Children*, 458-461. <https://doi.org/10.1145/2771839.2771885>.
- Peter, E. (2012). Critical thinking: Essence for teaching mathematics and mathematics problem solving skills. *African Journal of Mathematics and Computer Science Research*, 5. <https://doi.org/10.5897/ajmcsr11.161>
- Polit, D., & Beck, C. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health*, 29(5), 489-497. <https://doi.org/10.1002/nur.20147>
- Reikerås, E. (2020). Relations between play skills and mathematical skills in toddlers. *ZDM*, 52(4), 703-716. <https://doi.org/10.1007/s11858-020-01141-1>
- Sachdeva, S., & Eggen, P.-O. (2021). Learners' Critical Thinking About Learning Mathematics. *International Electronic Journal of Mathematics Education*, 16(1306-3030), em0644. <https://doi.org/10.29333/iejme/11003>
- Sinclair, N. (2018). Time, immersion and articulation: Digital technology for early childhood mathematics. In *Contemporary research and perspectives on early childhood mathematics education*. 205-221. Springer. https://doi.org/10.1007/978-3-319-73432-3_11
- Snyder, L., & Snyder, M. (2008). Teaching critical thinking and problem solving skills. *The Journal of Research in Business Education*, 50(2), 90.
- Strasser, J., & Bresson, L. (2017). *Big questions for young minds: Extending children's thinking*. National Association for the Education of Young Children.
- Sukstienwong, A. (2018). *Animo math: The role-playing game in mathematical learning for children*. *TEM Journal*, 7(1), 147-154. <https://doi.org/10.18421/TEM71-17>.
- Uğürel, I., & Morali, S. (2008). Matematik ve oyun etkileşimi [Interaction between mathematics and games]. *Journal of Gazi Educational Faculty [Gazi Eğitim Fakültesi Dergisi]*, 28(3), 75-98.