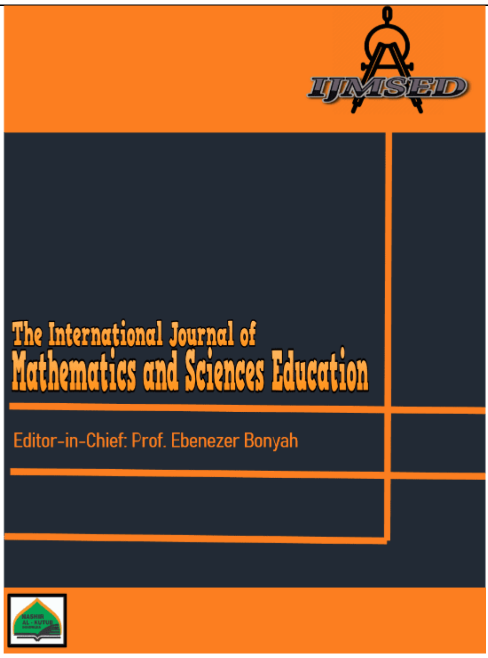
The International Journal of Mathematics and Sciences Education

Publication details, including instructions for authors and subscription information:
https://nakiscience.com/index.php/IJMSEd



Contextual learning integrated with gamification: Effects on mathematical connection ability and learning motivation in ratio and proportion

Nurmiati Nurmiati[1]*, Endang Wahyuningrum[2], Sudirman Sudirman[3], Tán Lì Nà[4]

[1]Department of Mathematics Education, Universitas Terbuka, Indonesia, nurmiati011@guru.smp.belajar.id

[2]Department of Mathematics Education, Universitas Terbuka, Indonesia, endagw@ecampus.ut.ac.id

[3]Department of Mathematics Education, Universitas Terbuka, Indonesia, Sudirman.official@ecampus.ut.ac.id

[4]Huanjiang High School, Guangxi, China, 1297164415@qq.com

To cite this article:
Nurmiati, N., Wahyuningrum, E., Sudirman, S., & Nà, T.. (2026). Contextual learning integrated with gamification: Effects on mathematical connection ability and learning motivation in ratio and proportion. *International Journal of Mathematics and Sciences Education*, *4*(1), 13-37

To link to this article:
https://nakiscience.com/index.php/IJMSEd

Published by:
Nasir Al-Kutub Indonesia
Residential Street Kila Rengganis, Block I, Number 11, Labuapi, Indonesia, 83361

International Journal of Mathematics and Sciences Education
Volume 4, Issue 1, pp. 13-37
https://nakiscience.com/index.php/IJMSEd

Contextualizing mathematics through local wisdom and gamification: Evidence from a quasi-experimental study on ratio and proportion

Nurmiati Nurmiati [1]*, Endang Wahyuningrum[2], Sudirman Sudirman[3], Tán Lì Nà[4]
[1]Department of Mathematics Education, Universitas Terbuka, Indonesia, nurmiati011@guru.smp.belajar.id
[2]Department of Mathematics Education, Universitas Terbuka, Indonesia, endagw@ecampus.ut.ac.id
[3]Department of Mathematics Education, Universitas Terbuka, Indonesia, Sudirman.official@ecampus.ut.ac.id
[4]Huanjiang High School, Guangxi, China, 1297164415@qq.com

*Corresponding author: nurmiati011@guru.smp.belajar.id

Mathematical connection ability remains a persistent weakness among Indonesian junior secondary students, with diagnostic data from SMP Negeri 2 Kaloran, Temanggung Regency, Indonesia showing that 74–81% of Grade 7 students failed to meet minimum competency on ratio and proportion across two academic years. Compounding this problem, classroom observations indicated chronically low learning motivation among the same cohort. This study examined whether contextual learning integrated with gamification (CTL+Gamification), delivered through the Wayground platform, could address both deficits simultaneously. A quasi-experimental design with nonequivalent control groups was employed, involving 56 Grade 7 students (experimental $n = 27$; control $n = 29$) selected through multi-stage cluster random sampling in Temanggung Regency. The experimental group received six sessions embedding ratio and proportion content in locally meaningful contexts, tobacco farming, coffee yield, and map scale, while the control group received conventional expository instruction. Mathematical connection ability was assessed using a validated seven-item essay test (Aiken's $V = 0.83$; $\alpha = 0.84$), and learning motivation using a 25-item ARCS-based questionnaire ($\alpha = 0.87$). Independent-sample t-tests and two-way ANOVA revealed that the experimental group significantly outperformed the control group in mathematical connection ability ($t = 3.542$, $p = .001$, $d = 0.97$) and learning motivation ($t = 2.201$, $p = .032$, $d = 0.59$), while the instructional model × motivational level interaction was nonsignificant ($F = 1.306$, $p = .280$). These findings suggest that contextually grounded, gamified instruction can strengthen both cognitive and affective outcomes, offering a locally adaptable model for mathematics teaching under Kurikulum Merdeka.

Article History
Received:
22 February 2026
Revised:
12 April 2026
Accepted:
24 April 2026
Published Online:
30 June 2026

Keywords:
Contextual learning; Gamification; Mathematical connection ability; Learning motivation; Ratio and proportion

1. INTRODUCTION

Mathematical connection ability, the capacity to recognize and apply relationships among mathematical concepts, across disciplines, and between mathematics and real-life contexts, is one of the core process standards outlined by the National Council of Teachers of Mathematics (NCTM, 2000). Over the past decade, research on this construct has matured considerably. Building on the framework originally introduced by Businskas (2008), the Extended Theory of Mathematical Connections has identified increasingly precise categories of connection-making that earlier models failed to capture, from the reasoning a teacher displays while explaining the derivative (Rodríguez-Nieto et al., 2022) to the finer distinctions revealed when this framework is combined with the onto-semiotic approach (Rodríguez-Nieto et al., 2023). Hatisaru (2024) reviews this trajectory and shows how the construct has grown from a single, general notion of "connection" into a differentiated set of analytical tools. Despite these theoretical advances, empirical studies consistently find that students struggle to make such connections in practice. Dolores-Flores et al. (2019) found that pre-university students frequently failed to link the concept of rate of change to its graphical and real-world representations, and García-García and Dolores-Flores (2021) similarly reported that students' intra-mathematical connections remained markedly weaker than their procedural fluency when solving calculus application problems. Indonesia's national curriculum reflects the same emphasis on connection-making: Keputusan Kepala BSKAP

The persistence of these learning difficulties is reflected not only in classroom-based studies but also in large-scale national and international assessments, which consistently reveal students' limited ability to apply mathematical knowledge in meaningful contexts. The 2022 PISA results illustrate the scale of the problem: only 18% of Indonesian students reached Level 2 mathematical proficiency, the threshold at which students begin applying mathematical concepts to contextual situations, compared with an OECD average of 69% (OECD, 2023). Fewer than 1% reached Level 5 or 6, where complex modeling and strategy evaluation are required. This gap between curriculum-based instruction and application-based mathematical literacy is not unique to Indonesia: Kang and Cogan (2022) found that even in systems with strong TIMSS-measured curriculum mastery, a large share of students still struggle to transfer that mastery into PISA-style contextual problem solving, particularly among lower-SES students. At the system level, Liu et al. (2024) show that the instructional practices most strongly associated with mathematics achievement across PISA cycles, cognitive activation and teacher support in particular, are precisely the practices that contextual, inquiry-based instruction is designed to strengthen, while Wang et al. (2023), synthesizing 156 studies on PISA mathematics predictors, identify socioeconomic status and opportunity to learn as the most consistent cross-national predictors of this same gap. The problem is not confined to low performers alone: Akmal and Pritchett (2021) demonstrate that even eliminating within-country rich–poor achievement gaps in developing countries would still leave most students far short of basic proficiency, and Gust et al. (2024) estimate that roughly two-thirds of the world's youth fail to reach basic skill thresholds on a common international scale, with the shortfall costing the global economy an estimated $700 trillion in lost output. Regional data tell a similar, if slightly more encouraging, story: the 2024 Asesmen Nasional numeracy score for Temanggung Regency stood at 68.12, an improvement from 66.48 the year before, though still only "intermediate" by provincial standards and far from reflecting adequate ability to apply mathematical concepts across real-world contexts.

Diagnostic data collected directly from SMP Negeri 2 Kaloran, a state junior secondary school in Kaloran District, Temanggung Regency, sharpen this picture further. As Table 1 shows, mean scores on ratio and proportion material sat at 50 in the 2023/2024 academic year and 53 the year before, with 81% and 74% of students respectively failing to reach the minimum competency threshold (KKM). Interviews with science and social studies teachers pointed to the same underlying weakness outside the mathematics classroom: students struggled to apply ratio concepts to numeracy tasks in IPA (measurement, pressure) and IPS (map scale, supply and demand), even though these applications require nothing beyond basic arithmetic.

Table 1
Mean Scores and Competency Attainment on Ratio and Proportion at SMP Negeri 2 Kaloran

Academic Year	Mean Score	% Reached KKM	% Below KKM
2023/2024	50	19%	81%
2022/2023	53	26%	74%

Low motivation appears to compound the problem. Teachers at SMP Negeri 2 Kaloran described students who disengaged from abstract mathematical content and rarely pushed beyond rote procedural steps, an observation that fits with Van Balen et al.'s (2024) finding of a significant positive relationship

International Journal of Mathematics and Sciences Education
Volume 4, Issue 1, pp. 13-37
https://nakiscience.com/index.php/IJMSEd

between motivation and mathematics achievement, and with Keller's (2010) ARCS model, which places These findings suggest that improving students' mathematical connection ability requires an instructional approach that simultaneously strengthens conceptual understanding and fosters sustained learning motivation.

Two instructional strategies speak directly to these twin problems of cognition and motivation. Contextual Teaching and Learning (CTL), as Johnson (2002) conceptualized it, anchors academic content in students' lived experience through seven components, constructivism, inquiry, questioning, learning community, modeling, reflection, and authentic assessment. In Indonesian classrooms specifically, CTL has repeatedly been shown to strengthen mathematical connection ability by grounding abstract concepts in familiar, personally meaningful situations (Muzaki et al., 2023; Kenedi et al., 2019). This aligns with broader international evidence: a recent meta-analysis of 28 studies found that contextualized approaches to mathematics instruction, including Realistic Mathematics Education, ethnomathematics, and technology integration, produce consistently strong positive effects on mathematical skills across educational levels and regions, with Realistic Mathematics Education showing the highest effectiveness among the approaches examined (Putri et al., 2025). More broadly, elementary mathematics programs that emphasize conceptual, connection-based instruction over rote procedure have also been shown to outperform conventional approaches in large-scale synthesis (Pellegrini et al., 2021). Temanggung Regency, an agricultural area built around coffee, tobacco, and vegetable farming, offers no shortage of authentic material for ratio and proportion, fertilizer composition, crop yield comparisons, map scale, temperature conversion all lend themselves naturally to CTL.

Gamification addresses the motivational side more directly. Defined by Deterding et al. (2011) and Wardani et al. (2025) as the use of game design elements, points, leaderboards, badges, timers, in non-game settings, it has become one of the more consistently effective tools for boosting learner engagement. A systematic review of 46 studies by Zainuddin et al. (2020) found reliable motivational and engagement benefits across secondary education, and closer to the present context. These findings converge with several large-scale meta-analyses: Sailer and Homner (2020) found significant effects of gamification on cognitive, motivational, and behavioral learning outcomes across the literature; Bai et al. (2020) similarly reported a medium positive effect on student learning outcomes across 30 independent interventions; and Huang et al. (2020) confirmed a comparable effect size across a broader corpus of gamification studies in formal educational settings. This pattern is consistent with the wider literature on educational technology in mathematics specifically, where meta-analyses have found modest but reliable positive effects on mathematics achievement (Li & Ma, 2010; Cheung & Slavin, 2013). This study uses the Wayground platform, which operationalizes gamification through real-time group leaderboards, tiered quiz challenges, badge awards, and immediate point feedback.

What remains unexamined is how these two approaches work together. CTL and gamification have each been studied on their own, but not in combination, and no prior work has asked whether their combined effect on mathematical connection ability and motivation depends on students' initial motivational level, nor has any study applied this combination to ratio and proportion material in Temanggung Regency using Wayground specifically. This gap matters theoretically as well as practically: motivation is not a uniform construct that responds identically to instruction regardless of a learner's starting point. A large-scale meta-analysis grounded in self-determination theory found that different types of motivation are associated with markedly different academic outcomes, suggesting that interventions may not benefit all learners equally depending on their initial motivational profile (Howard et al., 2021), and longitudinal evidence specifically in mathematics has shown that students' initial levels of motivation and amotivation differentially predict subsequent learning gains (Leroy & Bressoux, 2016). More directly still, recent aptitude-treatment-interaction research in mathematics education has demonstrated that instructional scaffolds and motivational supports do not benefit all students equally

1) Does mathematical connection ability differ significantly between students receiving CTL with gamification and those receiving conventional instruction?
2) Does learning motivation differ significantly between the two groups?
3) Does the instructional model interact significantly with students' motivational level in shaping mathematical connection ability?

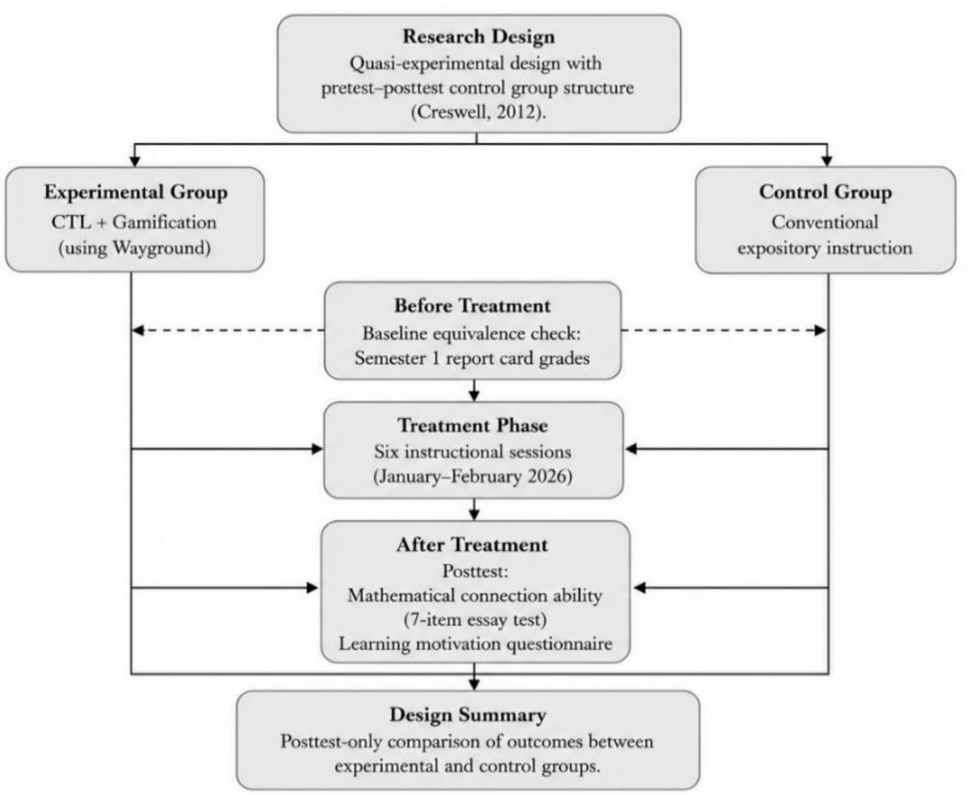
2. Methods

2.1 Research Design

This study employed a quasi-experimental design with a *pretest-posttest control group* structure (Creswell, 2012). This design was selected because random assignment of individual students to conditions was not feasible within intact school classrooms; instead, whole intact classes were assigned to treatment and control conditions, which is the defining feature of quasi-experimental educational research. The experimental group received contextual learning integrated with gamification (CTL+Gamification using Wayground), while the control group received conventional expository instruction on the same ratio and proportion content. Mathematical connection ability was measured at pretest (using semester 1 grades as baseline) and posttest (using a validated 7-item essay test).

Figure 1

Experimental research design



It should be noted that semester 1 report card grades were used solely to establish group equivalence prior to the intervention, not as a pretest measure of mathematical connection ability itself; no pretest using the target instrument was administered, as the 7-item essay test was developed and validated specifically for posttest use. This study therefore reports posttest-only group comparisons on the target outcome rather than a true pretest–posttest gain score on the same instrument; this is acknowledged as a limitation in Section 4. Learning motivation was measured only at posttest, after all sessions were completed, as motivation was treated as an outcome of the instructional treatment rather than a pre-existing covariate. Because motivation was assessed only at posttest and intact classes rather than individually randomized students were assigned to conditions, the observed between-group difference in motivation may partly reflect pre-existing differences between the two classes rather than solely the effect of the instructional treatment; this threat to internal validity is discussed further in Section 4 (See Figure 1).

The overall study procedure followed three sequential phases. In the preparation phase, the researcher conducted classroom observation, teacher interviews, and diagnostic data analysis; designed six lesson plans integrating CTL components with Wayground gamification; and developed and validated the test and questionnaire through expert review (Aiken's V) and a pilot study ($n = 30$). In the implementation phase (January–February 2026), the pretest was administered, six instructional sessions were conducted with an independent observer present, and the posttest and motivation questionnaire were administered immediately after the final session. In the analysis phase, data were scored, assumption tests were run, and statistical hypotheses were tested using independent-sample t-tests, two-way ANOVA, and Cohen's d effect sizes in IBM SPSS Statistics Version 25.

2.2 Participants and Research Setting

The study was conducted at SMP Negeri 2 Kaloran, Kaloran District, Temanggung Regency, Central Java, Indonesia, during the 2025/2026 academic year. Temanggung is a predominantly agricultural regency known for coffee, tobacco, and vegetable farming, providing authentic real-world contexts for learning ratio and proportion within the CTL approach.

Participants were selected using a multi-stage cluster random sampling procedure. First, one of the 20 sub-districts in Temanggung Regency (Kaloran District) was randomly selected. Next, the research school was randomly chosen from the three public junior secondary schools within the selected district. Finally, two intact Grade 7 classes were randomly assigned as the experimental and control groups. The experimental class consisted of 27 students, whereas the control class consisted of 29 students. Both groups were taught by the same mathematics teacher to minimize teacher-related variability during the intervention. Participant characteristics are summarized in Table 1.

Table 1
Participant Characteristics

Characteristic	Experimental Group	Control Group	Total
Grade Level	Grade 7	Grade 7	—
Number of Students	27	29	56
School	SMP Negeri 2 Kaloran	SMP Negeri 2 Kaloran	—
Academic Year	2025/2026	2025/2026	—
Instructional Model	CTL with Gamification	Conventional Instruction	—
Mathematics Teacher	Same teacher	Same teacher	—

Participant eligibility required active enrollment in Grade 7 during the 2025/2026 academic year. All students in the selected classes were included in the study, and no minimum attendance threshold was imposed for inclusion in the analysis. Baseline equivalence between groups was examined using an independent-samples t-test based on first-semester mathematics grades as a proxy for prior mathematical achievement. The analysis indicated no statistically significant difference between the experimental and control groups ($p > .05$), supporting the comparability of the groups before the intervention.

The study received formal approval from the school principal. Written informed consent was obtained from parents or legal guardians, informed consent was obtained from the participating mathematics teacher, and verbal assent was obtained from all students prior to data collection. All procedures complied with ethical standards for educational research involving minors.

Because intact classes rather than individually randomized students were assigned to the intervention conditions, classroom-level characteristics such as peer interaction and classroom culture could not be fully controlled. Likewise, although the same teacher taught both groups, the possibility of a novelty or Hawthorne effect cannot be entirely excluded. Furthermore, the sample size (27 and 29 students) reflected the enrollment of the available intact classes rather than the result of an a priori statistical power analysis. These methodological limitations are discussed further in Section 4.

2.3 Instructional Intervention

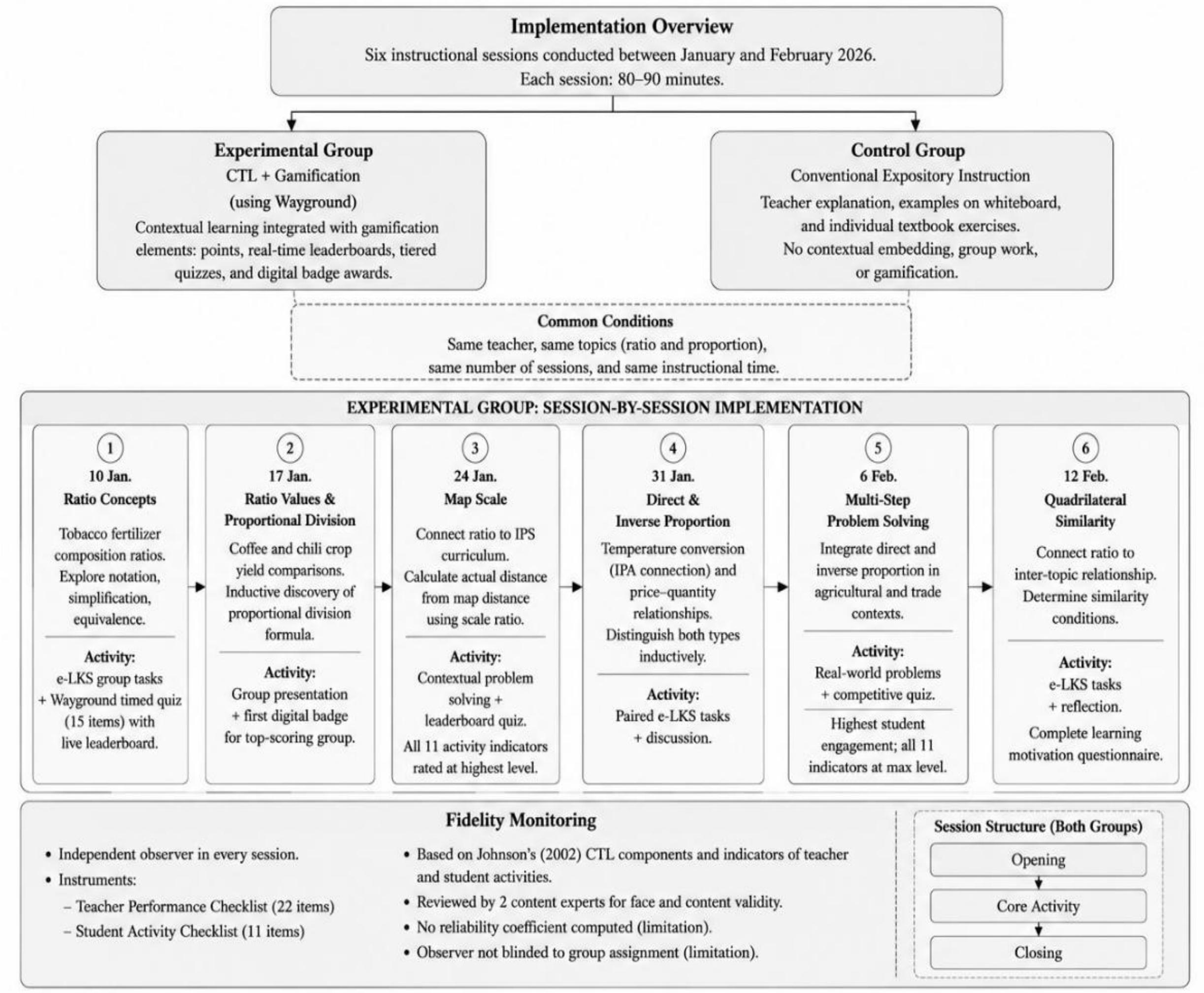
The experimental group received six sessions of CTL+Gamification conducted between January and February 2026, each lasting 80–90 minutes. Every session followed a consistent three-phase structure, Opening, Core Activity, and Closing, within which the CTL components were systematically embedded alongside Wayground's gamification elements, including points, real-time leaderboards, tiered quizzes, and digital badge awards. Students worked in heterogeneous groups of three to four members and shared at least one device to access Wayground.

The teacher acted as a facilitator throughout the intervention, while an independent observer recorded implementation fidelity using a structured 22-item checklist for teacher performance and an 11-item checklist for student activity in every session. The teacher performance and student activity checklists were developed based on Johnson's (2002) Contextual Teaching and Learning framework, the planned instructional stages presented in the lesson plans, and indicators of teacher and student activities during CTL-based learning. The instruments were reviewed by two content experts in mathematics education and educational evaluation to assess their clarity, relevance, and alignment with the instructional treatment. However, no formal reliability coefficient was computed for these fidelity instruments, which is acknowledged as a limitation in Section 4. Because the two instructional conditions differed visibly in their structure and learning materials, the independent observer could not be blinded to group assignment. This condition may have introduced observer expectancy bias into the implementation-fidelity and student-engagement ratings and is therefore acknowledged as a study limitation.

The six intervention sessions were conducted as follows. Session 1, held on January 10, addressed ratio concepts through the context of tobacco-fertilizer composition. Students explored ratio notation, simplification, and equivalence using group e-LKS activities, followed by a 15-item timed Wayground quiz with a live leaderboard. Session 2, held on January 17, addressed ratio values and proportional division through comparisons of coffee and chili crop yields. Students worked collaboratively to derive the proportional-division formula and presented their solutions, while the highest-scoring group received a digital badge. Session 3, held on January 24, focused on map scale and connected ratio concepts with the social studies curriculum. Students calculated actual distances from map distances using scale ratios. Session 4, held on January 31, addressed direct and inverse proportion through temperature conversion and price–quantity relationships. Students distinguished the two types of proportion through paired e-LKS activities. Session 5, held on February 6, involved multi-step problem solving that integrated direct and inverse proportion in agricultural and trading contexts. Session 6, held on February 12, addressed quadrilateral similarity and connected ratio with relationships among mathematical topics. Students determined similarity conditions through e-LKS activities and completed the learning-motivation questionnaire at the end of the session (see Figure 2).

Figure 2

Implementation of Instructional Treatment



The control group received six parallel sessions of conventional expository instruction on the same ratio and proportion topics. The sessions were taught by the same teacher and involved concept explanation, worked examples on the whiteboard, and individual textbook exercises, without contextual learning activities, structured group work, or gamification. The use of the same teacher, content coverage, number of sessions, and instructional duration was intended to reduce differences attributable to teacher characteristics, content exposure, and instructional time.

2.4 Instruments

Two instruments were employed to collect the research data: a Mathematical Connection Ability Test and a Learning Motivation Questionnaire. Prior to implementation, both instruments underwent expert validation and pilot testing to establish their validity and reliability. The main characteristics of the instruments are summarized in Table 2.

International Journal of Mathematics and Sciences Education
Volume 4, Issue 1, pp. 13-37
https://nakiscience.com/index.php/IJMSEd

Table 2
Research Instruments

Aspect	Mathematical Connection Ability Test	Learning Motivation Questionnaire
Purpose	Measure students' mathematical connection ability	Measure students' learning motivation
Construct	NCTM (2000) Mathematical Connections	Keller's (2010) ARCS Model
Number of Items	7 essay questions	25 Likert-scale items
Scale	Rubric (0–4 per item)	4-point Likert scale (1 = Strongly Disagree to 4 = Strongly Agree)
Maximum Score	28 (converted to 0–100)	100
Indicators	(1) Connections among mathematical topics; (2) Connections with other disciplines; (3) Connections with real-life contexts	Attention (6), Relevance (7), Confidence (6), Satisfaction (6)
Content Validity	Aiken's V = 0.83 (all items > 0.75)	Expert validation
Reliability	Cronbach's α = 0.84	Cronbach's α = 0.87
Administration	Individual posttest after Session 6	Individual questionnaire after Session 6

The Mathematical Connection Ability Test consisted of seven essay questions designed to assess the three mathematical connection indicators proposed by the National Council of Teachers of Mathematics (NCTM, 2000). The items represented connections among mathematical topics, connections between mathematics and other disciplines, and connections between mathematics and real-life situations within the topic of ratio and proportion. Students' responses were scored using a four-point analytic rubric (0–4). Raw scores were converted to a 0–100 scale using the formula: Converted Score = (Raw Score ÷ 28) × 100 to facilitate comparison with learning motivation scores. Essay responses were evaluated by one trained scorer using the standardized scoring rubric.

The Learning Motivation Questionnaire was developed based on Keller's (2010) ARCS motivational model and adapted from Sardiman (2014). The questionnaire comprised 25 statements covering the four motivational dimensions: Attention, Relevance, Confidence, and Satisfaction. Students responded using a four-point Likert scale ranging from 1 (Strongly Disagree) to 4 (Strongly Agree). The questionnaire was administered individually after completion of the final learning session to avoid influencing students' learning behavior during the intervention. To facilitate subsequent statistical analyses, each student's test and questionnaire responses were linked using confidential identification codes accessible only to the research team. Individual responses remained confidential and were not disclosed to teachers or other students.

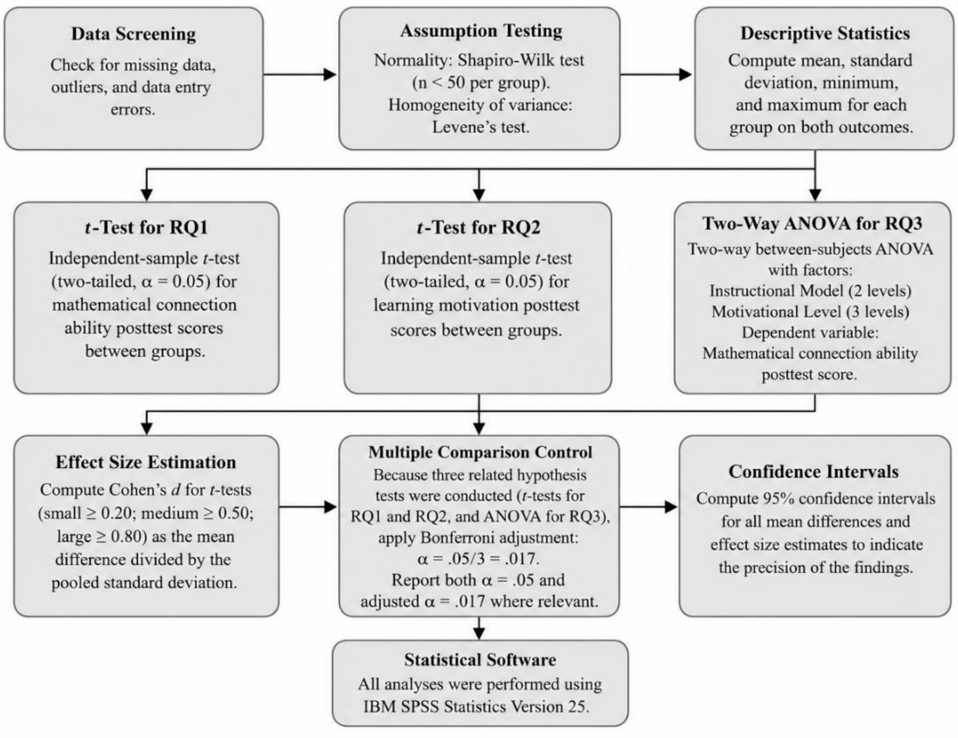
2.5 Data Analysis

Prior to hypothesis testing, statistical assumptions were verified. Normality of posttest score distributions was assessed using the Shapiro-Wilk test, recommended for small samples ($n < 50$ per group). Homogeneity of variance was assessed using Levene's test. Both tests were applied to mathematical connection ability and learning motivation scores. Descriptive statistics (mean, SD, minimum, maximum) were computed for both groups on both outcome variables. For Research Questions 1 and 2, independent-sample t-tests were conducted at $\alpha = 0.05$ (two-tailed). Practical significance was quantified using Cohen's d (small ≥ 0.20; medium ≥ 0.50; large ≥ 0.80; Cohen, 1988), computed as the mean difference divided by the pooled standard deviation. Because three related hypothesis tests were conducted across the study's research questions (the t-tests for RQ1 and RQ2, and the ANOVA for RQ3), the family-wise Type I error rate was considered; a Bonferroni-adjusted threshold (α = .05/3 = .017) is reported alongside the conventional threshold in Section 3 where relevant. In addition, 95% confidence intervals were computed for all mean differences and effect size estimates to provide additional information on the precision of the findings.

For Research Question 3, a two-way between-subjects ANOVA was conducted with instructional model (CTL+Gamification vs. conventional) and motivational level as fixed between-subjects factors, and mathematical connection ability posttest score as the dependent variable. The A×B interaction effect was the primary focus for Research Question 3, tested at $\alpha = 0.05$. A non-significant interaction ($p > 0.05$) indicates that the instructional model advantage is consistent across all motivational subgroups, which has important implications for inclusive classroom practice. All analyses were performed using IBM SPSS Statistics Version 25 (See Figure 3).

Figure 3

Data Analysis Process



3. Results

3.1 Implementation of CTL with Gamification

The contextual teaching and learning (CTL) intervention integrated with Wayground gamification was implemented in six meetings from January to February 2026. Each meeting lasted approximately 80–90 minutes and followed the same instructional sequence: opening, core learning activities, and closing. The experimental class received contextual problem situations, collaborative electronic student worksheets (e-LKS), group discussion, presentation, timed Wayground quizzes, real-time rankings, and digital rewards, while the control class studied the same ratio and proportion content for the same duration through conventional expository instruction. The photographic and documentary evidence below illustrates how this design was translated into classroom practice across four connected stages: contextual problem orientation, whole-class orientation, collaborative investigation, and gamified practic.

Each meeting opened with contextual problem orientation. Figure 4 presents a comparison task that introduces ratio through the heights of two school plants, an Aglaonema measuring 60 cm and a palm tree measuring 2 m, requiring students to convert units before determining the ratio between the two heights. Rather than presenting ratio as an abstract procedure, the task operationalizes the CTL principle of meaningful connection: students link mathematical symbols to observable objects and must explain why quantities in different units cannot be compared directly before conversion. This concrete entry point was supported by the e-LKS shown in Figure 5, which communicates the learning objectives, provides navigation instructions, and opens with discussion questions rather than direct formulas, activating prior knowledge and generating cognitive conflict before formal explanation. Because the worksheet was accessed digitally and followed a consistent format across meetings, contextual situation, strategy discussion, response entry, feedback, next task, students could move fluidly between explanation, examples, and Wayground activities, and procedural confusion decreased as the sessions progressed.

Figure 4

A contextual comparison task used to introduce ratio through the heights of two school plants

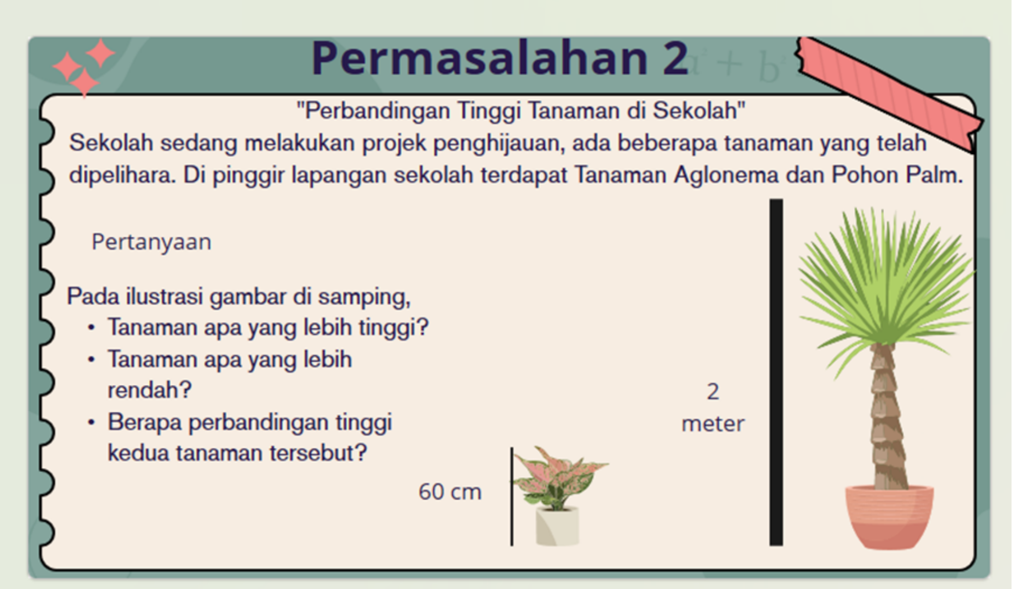
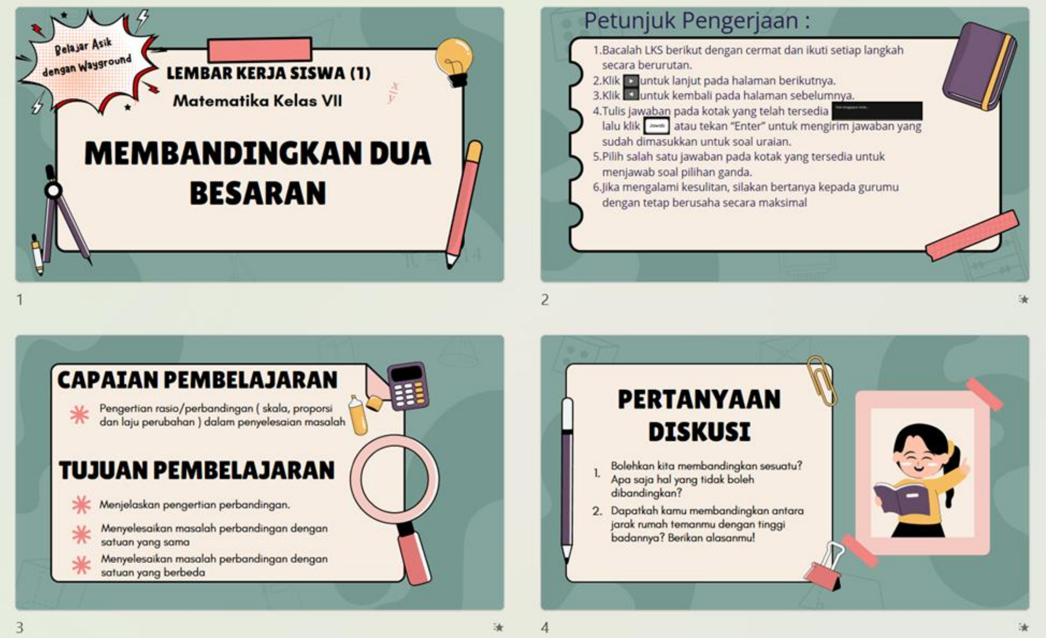


Figure 5

Overview of the electronic student worksheet, including instructions, learning objectives, and initial discussion prompts



International Journal of Mathematics and Sciences Education
Volume 4, Issue 1, pp. 13-37
https://nakiscience.com/index.php/IJMSEd

Whole-class orientation, shown in Figure 6, preceded group work in every session. The teacher introduced the learning goals, clarified the contextual problem, and explained the sequence of collaborative and digital activities, asking students to predict possible solutions and identify the information needed to solve the problem. This stage functioned as inquiry preparation rather than lecture: the teacher acted as facilitator, posing guiding questions and confirming procedural instructions without disclosing the final formula, which helped keep treatment delivery consistent across the six sessions.

Figure 6

Whole-class orientation before students began collaborative CTL and Wayground activities

Collaborative investigation followed in heterogeneous groups of three to four students, as documented in Figures 7 and 8. Groups discussed the contextual task, completed the e-LKS, and shared at least one device, dividing responsibilities, checking calculations, and negotiating a final group response. The teacher circulated among groups, asked probing questions, for example, requesting justification of a unit conversion, comparison of two solution strategies, or explanation of whether a relationship was directly or inversely proportional, and provided scaffolding where needed, reflecting the CTL components of learning community, inquiry, questioning, and constructivism. Figure 8 provides a closer view of this activity: students recorded calculations in writing while using the shared mobile device to access tasks and verify responses, confirming that Wayground served as an engagement and feedback tool rather than a replacement for mathematical reasoning. Students were also observed responding to one another's work, indicating genuine peer communication rather than isolated device use; across sessions, the observer noted increasingly efficient group coordination and more frequent student-initiated explanations.

International Journal of Mathematics and Sciences Education
Volume 4, Issue 1, pp. 13-37
https://nakiscience.com/index.php/IJMSEd

Figure 7
Students working in heterogeneous groups while the teacher facilitates discussion and monitor's progress

Figure 8
Active group problem solving using written work and a shared mobile device

Each meeting closed with gamified practice, shown in Figure 9. After completing the contextual investigation, students answered timed Wayground questions on the same concept, with points, real-time rankings, tiered difficulty, and digital badges providing immediate performance feedback. These elements sustained attention and encouraged repeated practice, while the teacher reminded students that accuracy and explanation mattered more than speed; students discussed each item before submitting a response, particularly when a question required unit conversion or a choice between direct and inverse proportion. Competition was organized at the group level to preserve collaboration despite the visible excitement generated by the live leaderboard. Student reflections indicated that the time limit encouraged concentration, the ranking system made lessons more enjoyable, and the contextual examples made ratio and proportion easier to understand.

Figure 9

Students completing Wayground activities collaboratively using shared mobile devices

Across all six sessions, implementation fidelity was recorded by an independent observer, RN using a 22-item teacher-performance checklist and an 11-item student-activity checklist in every meeting. All 22 teacher-performance indicators were implemented in Sessions 1–5; in Session 6, 21 of the 22 indicators were implemented, with the planned whole-class reflection and summary shortened because students needed additional time to complete the learning motivation questionnaire after the final instructional task. The teacher nevertheless gave a brief oral conclusion and clarified the main concept before the questionnaire was distributed. Because the omitted activity occurred after the core CTL, group-work, and Wayground components had already been completed, this deviation is judged minor, though it is reported transparently as a small reduction in fidelity during the final session.

Student-activity ratings showed a progressive improvement across the intervention. In Session 1, several indicators were rated at Level 3, most, but not all, students participating actively, a pattern associated with students' initial adjustment to the digital worksheet, shared-device arrangement, group roles, and timed quizzes. Participation became more consistent as students grew familiar with the routine, and by Session 5 all 11 student-activity indicators were rated at Level 4, indicating participation by all students, with the strongest engagement observed during contextual problem solving, group discussion, response submission, and reaction to immediate Wayground feedback. This progression suggests that engagement was not merely an initial novelty response but reflected increasingly stable participation routines built through repeated exposure to the intervention.

Student reflections submitted through Wayground were consistent with the observation results. In early meetings, students described mathematics as more enjoyable and the questions as challenging; in later sessions their responses became more specific, reporting that agricultural, trade, temperature, and map-scale contexts helped them understand why ratio and proportion were useful. Several students noted that the Temanggung map task made scale easier to understand because the calculation was tied to an identifiable place, and that the leaderboard encouraged focus while digital badges created a sense of group achievement. By the final session, many students requested that Wayground and collaborative contextual tasks continue to be used in future mathematics lessons.

Taken together, the photographs, observation checklists, and student reflections provide convergent process evidence that the intervention was delivered substantially as planned, demonstrating the presence of its essential features: contextual problem orientation, student inquiry, collaborative learning, teacher facilitation, written mathematical reasoning, gamified practice, immediate feedback, and reflection. Using the same teacher, equivalent instructional time, and identical content coverage in the experimental and control classes reduced alternative explanations related to teacher quality or exposure to content. The observer could not, however, be blinded to group assignment because the two instructional conditions were visibly different, so some observer-expectancy influence on fidelity and engagement ratings cannot be ruled out. Despite this limitation, the high consistency of teacher implementation, the progressive student-activity ratings, and the alignment between observation and student reflection support the conclusion that the CTL-with-gamification treatment was implemented.

3.2 Mathematical Connection Ability

To provide an initial overview of the intervention outcomes, descriptive statistical analysis was conducted for both the experimental and control groups. The analysis summarizes students' mathematical connection ability before and after the implementation of the CTL with gamification model, including measures of central tendency and score distribution. Table 3 presents descriptive statistics for mathematical connection ability in both groups at pretest and posttest.

Table 3

Descriptive Statistics for Mathematical Connection Ability

Statistic	Pretest Exp.	Pretest Ctrl.	Posttest Exp.	Posttest Ctrl.
Highest Score	54	64	91	81
Lowest Score	31	26	34	20
Mean ($\bar{x}$)	43	39	65	48
SD	6.16	8.91	17.41	17.64
n	27	29	27	29

Table 3 reveals that both groups were comparable at pretest (experimental: $\bar{x} = 43$; control: $\bar{x} = 39$; difference = 4 points), confirming baseline equivalence. After the six-session intervention, the experimental group's posttest mean ($\bar{x} = 65$) exceeded the control group ($\bar{x} = 48$) by 17 points. The experimental group gained 22 points from pretest to posttest, more than double the control group's 9-point gain. The nearly identical standard deviations at posttest (17.41 vs. 17.64) indicate that the intervention benefited students across the full ability spectrum within the experimental class, not only the highest achievers, suggesting the inclusive effectiveness of the CTL+Gamification model.

Prior to hypothesis testing, normality and homogeneity of posttest scores were verified. Shapiro-Wilk tests confirmed normal distribution for both groups (experimental: $W = 0.932$, $df = 27$, $p = 0.078 > 0.05$; control: $W = 0.953$, $df = 29$, $p = 0.248 > 0.05$). Levene's test confirmed homogeneity of variance (Levene statistic = 0.298, $df1 = 1$, $df_2 = 54$, $p = 0.587 > 0.05$). The parametric t-test assumptions were therefore met.

Table 4

Hypothesis Test Results for Mathematical Connection Ability (Independent-Sample t-Test)

Variable	t-count	t-table	Sig. (2-tailed)	Cohen's d	Decision
Mathematical Connection Ability	3.542	2.003	0.001	0.97 (Large)	H_0 Rejected

Table 4 shows t-count (3.542) > t-table (2.003) with Sig. (2-tailed) = 0.001 < 0.05. H_0 is rejected: CTL with gamification had a significant positive effect on mathematical connection ability. Cohen's d = 0.97 indicates a large effect (Cohen, 1988), meaning the average experimental student exceeded approximately 83% of the control group in mathematical connection ability. This represents a practically meaningful improvement beyond statistical significance alone.

3.3 Learning Motivation

To obtain a broader understanding of students' affective responses to the intervention, descriptive analysis of learning motivation was conducted for both the experimental and control groups. The analysis describes the distribution of students across motivation categories and compares the overall motivation levels after the implementation of the CTL with gamification model. Table 5 presents the distribution of learning motivation categories in both groups after the intervention, based on Keller's (2010) ARCS model measured by the validated questionnaire (α = 0.87).

Table 5

Distribution of Learning Motivation Categories by Group

Category	Score Range	Exp. (n)	Exp. (%)	Ctrl. (n)	Ctrl. (%)
High	≥ 83.9	6	22%	4	14%
Moderate	66.7 – 83.9	19	70%	20	69%
Low	≤ 66.7	2	8%	5	17%
Total	–	27	100%	29	100%
Mean ($\bar{x}$)	–	77.63	SD = 9.22	72.90	SD = 6.77

Table 5 shows clear distributional differences favoring the experimental group: the proportion of high-motivation students was higher (22% vs. 14%), the proportion of low-motivation students was lower (8% vs. 17%), and the mean motivation score was higher (77.63 vs. 72.90). These differences are substantively meaningful: the proportion of students with low motivation was cut by more than half in the experimental group compared to the control, indicating that the CTL+Gamification model was particularly effective in elevating initially disengaged students. Normality (experimental: p = 0.442; control: p = 0.296) and homogeneity (Levene p = 0.097) assumptions were satisfied.

Table 6

Hypothesis Test Results for Learning Motivation (Independent-Sample t-Test)

Variable	t-count	t-table	Sig. (2-tailed)	Cohen's d	Decision
Learning Motivation	2.201	2.003	0.032	0.59 (Moderate)	H_0 Rejected

Table 6 shows t-count (2.201) > t-table (2.003) with Sig. (2-tailed) = 0.032 < 0.05. H_0 is rejected: CTL with gamification produced a significant positive effect on learning motivation. Cohen's d = 0.59 indicates a moderate effect, practically meaningful and exceeding the d = 0.50 threshold. The smaller effect size for motivation compared to connection ability (d = 0.97) reflects the theoretically expected differential responsiveness of affective versus cognitive constructs to short-term instructional interventions (Pintrich & Schunk, 2002).

3.4 Interaction Between Instructional Model and Motivational Level

To examine whether the effectiveness of the instructional model differed according to students' motivational level, descriptive statistics were first compared across motivational categories. Subsequently, a two-way ANOVA was conducted to determine the main effects of instructional model and motivational level, as well as their interaction on students' mathematical connection ability.

Table 7

Mathematical Connection Ability by Group and Motivational Level

Motivational Level	Exp. Mean	Exp. n	Ctrl. Mean	Ctrl. n	Difference
High	78.88	8	56.00	4	22.88
Moderate	61.88	17	54.05	19	7.83
Low	35.00	2	25.17	6	9.83
Total	65.00	27	48.00	29	17.00

Table 7 indicates that students in the experimental group consistently achieved higher mathematical connection scores than those in the control group across all motivational levels. The largest mean difference was observed among students with high motivation (78.88 vs. 56.00; difference = 22.88 points), followed by students with low motivation (35.00 vs. 25.17; difference = 9.83 points) and moderate motivation (61.88 vs. 54.05; difference = 7.83 points). Overall, the experimental group obtained a higher average score (65.00) than the control group (48.00), representing an overall advantage of 17 points. These descriptive findings suggest that the CTL with gamification model improved mathematical connection ability regardless of students' initial motivational level.

Table 8

Two-Way ANOVA Results for Mathematical Connection Ability

Source of Variation	Type III SS	df	Mean Square	F	Sig.
Corrected Model	11,426.133	5	2,285.227	12.752	0.000
Intercept	83,864.931	1	83,864.931	467.974	0.000
Instructional Model (A)	1,425.117	1	1,425.117	7.952	0.007
Motivational Level (B)	5,538.138	2	2,769.069	15.452	0.000
A × B Interaction	468.095	2	234.048	1.306	0.280
Error	8,960.420	50	179.208	–	–

Table 8 shows significant main effects of instructional model ($F(1,50) = 7.952$, $p = 0.007$, $\eta^2 p = 0.137$) and motivational level ($F(2,50) = 15.452$, $p < 0.001$, $\eta^2 p = 0.382$). According to Cohen's (1988) benchmarks, the instructional model demonstrated a medium-to-large effect, while motivational level exhibited a large effect on students' mathematical connection ability. These findings indicate that both the learning approach and students' motivational level independently contributed to differences in mathematical connection performance.

In contrast, the interaction between instructional model and motivational level was not statistically significant ($F(2,50) = 1.306$, $p = 0.280$, $\eta^2 p = 0.050$), indicating a small interaction effect. This result suggests that the effectiveness of the CTL with gamification model did not differ significantly across students with high, moderate, and low levels of motivation. In other words, the instructional model produced relatively consistent improvements regardless of motivational category.

International Journal of Mathematics and Sciences Education
Volume 4, Issue 1, pp. 13-37
https://nakiscience.com/index.php/IJMSEd

The descriptive results presented in Table 6 support this conclusion. Students in the experimental group achieved higher mean scores than those in the control group at every motivational level, with advantages of 22.88 points for highly motivated students, 7.83 points for moderately motivated students, and 9.83 points for students with low motivation. The absence of a crossover pattern among the group means is consistent with the non-significant interaction observed in the two-way ANOVA, indicating that the superiority of the CTL with gamification model remained relatively stable across motivational groups.

4. Discussion

The large effect size for mathematical connection ability (Cohen's d = 0.97) is explained by the systematic alignment between the CTL+Gamification design and the three NCTM (2000) connection indicators. The six contextual sessions addressed each indicator directly: Sessions 1–2 (tobacco fertilizer, coffee yield) embedded ratio concepts in real-world agricultural contexts familiar to students in Temanggung, addressing the third indicator (connections to everyday life); Sessions 3–4 (map scale and temperature conversion) explicitly connected ratio to IPS (Temanggung–Wonosobo map distance) and IPA (Celsius–Réaumur conversion), addressing the second indicator (cross-disciplinary connections); and Session 6 (photo frame similarity/kesebangunan) connected ratio concepts to a related mathematical topic, addressing the first indicator (connections among mathematical topics). This systematic coverage across all three NCTM indicators distinguishes the intervention from conventional instruction, which typically addresses ratio procedurally without contextual embedding. This approach is consistent with a narrative synthesis of pedagogical models for contextual mathematics learning, which identifies cross-disciplinary linkage and authentic situational grounding as the primary drivers of CTL's long-term effectiveness (Adeoye, 2024). These findings also reinforce evidence that mathematical connections built through real contexts can be analyzed through an onto-semiotic lens across application domains, as demonstrated in engineering students' mathematical connections in vector problem-solving (Rodríguez-Nieto et al., 2024).

The gamification layer through Wayground reinforced these connections through multiple complementary mechanisms. The time-limited quizzes created productive urgency that required students to retrieve and apply prior knowledge efficiently, consistent with the desirable difficulties framework (Bjork and Bjork, 2011). This effect has also been confirmed in advanced mathematics learning contexts, where spaced retrieval practice on mathematics quizzes was shown to reduce short-term performance while significantly increasing long-term retention (Lyle et al., 2022). Immediate point feedback after each formative quiz (an average of 18 items per session) provided the kind of timely corrective information that Hattie and Timperley (2007) identified as among the most powerful learning influences. The effectiveness of this type of formative feedback in mathematics learning specifically is further corroborated by an experimental study showing that structured formative assessment practice improves student achievement (Boström and Palm, 2023), and by a systematic literature review mapping which feedback characteristics are most influential in mathematics education research (Söderström and Palm, 2024). The group leaderboard created collaborative accountability in which every group member was motivated to understand the material rather than free-ride, operationalizing Vygotsky's (1978) Zone of Proximal Development through peer-assisted collaborative learning. This leaderboard-driven collaborative accountability mechanism aligns with a recent systematic review of leaderboard design in higher education (Li et al., 2024), and with findings that cognitively-guided peer tutoring patterns produce significantly higher group achievement in mathematical problem-solving (Li, Jiang, et al., 2023). These findings align with Kenedi et al. (2019), who reported d = 0.84 for contextual approaches on mathematical connection ability across 25 studies, with Muzaki et al. (2023), and with a large-scale meta-analysis confirming a significant, large overall effect of gamification on teaching and learning

International Journal of Mathematics and Sciences Education
Volume 4, Issue 1, pp. 13-37
https://nakiscience.com/index.php/IJMSEd

The moderate effect on learning motivation (d = 0.59) is theoretically explained through the four ARCS components (Keller, 2010; Gyenyn et al., 2026) as operationalized in the intervention. This ARCS framework has also proven effective in other gamified digital game-based learning contexts, where all four dimensions contributed significantly to increased motivation and academic performance (Camacho-Sánchez et al., 2023). Attention was captured through Wayground's visual interface, real-time leaderboard displays, and quiz timers. Relevance was established through all six Temanggung-contextualized scenarios, when students encountered ratio problems about coffee crop yields or tobacco fertilizer needs, they experienced the direct personal relevance that Ryan and Deci (2000) identify as essential for internalized motivation. Confidence was built through tiered quiz difficulty, providing the graduated mastery experiences that Bandura (1997) identifies as the most powerful source of academic self-efficacy. This effect of tiered tasks on mathematical self-efficacy is further confirmed experimentally, where students working on mathematics tasks marked with difficulty levels (easy–medium–hard) reported higher self-efficacy than those without such markers (Herset et al., 2023); a comparable pattern emerged in leaderboard-based gamification in a calculus course, although the mediating role of self-efficacy on learning performance required more careful gamification design (Ortiz-Rojas et al., 2025). Satisfaction was generated through point accumulation, badge awards, and group recognition — mechanisms that Deci and Ryan (1985) describe as bridging extrinsic and intrinsic motivation. This relationship is reinforced by a recent meta-analysis showing that gamification enhances students' intrinsic motivation and perceptions of autonomy and relatedness, although its impact on perceived competence remains minimal (Li, et al, 2024). The shift of low-motivation students from 17% (control) to 8% (experimental) is consistent with Luarn et al.'s (2023) finding that gamification is particularly impactful for initially low-motivation learners.

The non-significant interaction effect (F = 1.306, p = 0.280) has important practical implications. It confirms that the effectiveness of CTL+Gamification was not contingent on students' pre-existing motivational characteristics; the model produced consistent benefits across high, moderate, and low motivation subgroups. This principle of main effect dominance (Shadish et al., 2002) means that mathematics teachers in Temanggung and comparable Indonesian contexts can implement this integrated approach across heterogeneous classes without pre-assessing or sorting students by motivational level. This inclusive effectiveness aligns with Rose and Meyer's (2002) Universal Design for Learning principle of multiple means of engagement, and with Tomlinson's (2001) criterion that effective differentiated instruction serves all learners. Empirical support for this UDL principle is further reinforced by a systematic review and meta-analysis demonstrating the effectiveness of UDL implementation in supporting student diversity within inclusive educational settings (Almeqdad et al., 2023). This inclusive effectiveness also directly supports the equity goals embedded in Indonesia's Kurikulum Merdeka, which emphasizes differentiated, student-centered learning that is accessible to students with diverse readiness profiles.

The study also contributes to the broader theoretical understanding of how CTL and gamification interact when combined. The present findings extend Ningsih (2021) by demonstrating that CTL+Gamification produces large effects not only on conceptual understanding but specifically on mathematical connection ability, a higher-order process standard that requires students to recognize and apply relationships across multiple representational and contextual domains. The findings also extend Freudenthal's (1991) Realistic Mathematics Education (RME) principle that mathematics learned as human activity in real contexts produces more transferable understanding, while demonstrating that

Several limitations warrant acknowledgment. First, the relatively small sample (n = 56 total) from a single school limits external validity; replication across multiple schools and districts is needed to strengthen generalizability. Second, the six-session timeframe does not permit assessment of the durability of motivational gains; longitudinal designs are needed to examine whether gamification-

International Journal of Mathematics and Sciences Education
Volume 4, Issue 1, pp. 13-37
https://nakiscience.com/index.php/IJMSEd

induced motivation is sustained after the intervention ends. Third, the present design cannot disentangle the independent contributions of CTL versus gamification to the observed outcomes; factorial designs with CTL-only, gamification-only, and combined conditions would provide more granular evidence. Fourth, the low-motivation subgroup in the experimental condition comprised only two students, making interaction estimates for that cell statistically imprecise. Fifth, baseline equivalence between groups was established using general mathematics achievement (semester grades) rather than a pretest of mathematical connection ability itself, so equivalence on the specific target construct could not be directly confirmed. Sixth, because intact classes rather than individually randomized students were assigned to conditions, class-level differences beyond semester grades (e.g., classroom culture, peer dynamics) and a possible teacher novelty/Hawthorne effect favoring the newly introduced gamified condition cannot be ruled out. Seventh, learning motivation was measured only at posttest without a baseline assessment, so the observed between-group difference may partly reflect pre-existing differences between the two classes rather than solely the treatment effect. Eighth, the independent observer could not be blinded to group condition, introducing a possible observer expectancy bias into fidelity and engagement ratings. Ninth, the implementation fidelity checklists, unlike the two main outcome instruments, did not undergo formal validity or reliability testing. Finally, the non-significant instructional model × motivational level interaction should be interpreted cautiously given the very small low-motivation experimental subgroup ($n = 2$), which limited statistical power to detect a true interaction effect; this finding is therefore better characterized as inconclusive rather than as confirmatory evidence of uniformly consistent effectiveness across motivational levels.

5. Conclusions

This study set out to determine whether CTL integrated with Wayground gamification produces greater gains in mathematical connection ability and learning motivation than conventional expository instruction, and whether this effect depends on students' initial motivational level. The findings answer the first two questions affirmatively and the third negatively, closing the gap identified in the introduction: no prior study had simultaneously examined the combined effect of CTL and gamification on both a cognitive and an affective outcome in the context of ratio and proportion instruction in Temanggung Regency.

For mathematical connection ability, the experimental group achieved a substantially higher posttest performance than the control group, with a large practical effect. For learning motivation, the experimental group also outperformed the control group, with a marked reduction in the proportion of low-motivation students, suggesting that the affective benefits of gamification extended even to initially disengaged learners. The interaction between instructional model and motivational level on mathematical connection ability was not found to be statistically detectable; given the study's sample size, this result should be read as no interaction detected rather than confirmed evidence of independence, and it warrants replication with adequately powered samples before being treated as a stable finding. Descriptively, however, gains were observed consistently across high-, moderate-, and low-motivation subgroups alike, a pattern consistent with, though not proof of, a broadly distributed cognitive benefit.

These findings contribute to the mathematics education literature in two ways. Theoretically, this is the first study to provide empirical evidence, specific to ratio and proportion instruction, that combining CTL with gamification can advance a cognitive outcome (mathematical connection ability) and an affective outcome (learning motivation) within the same intervention, addressing a gap left by prior studies that examined CTL and gamification separately. Practically, the study demonstrates that this dual benefit is achievable using a freely accessible gamification platform and locally contextualized materials, without additional infrastructure, offering classroom teachers a concrete, low-cost instructional sequence rather than a general recommendation to "use gamification." For curriculum developers and school leaders, the findings suggest that professional development on contextual-gamified design may be a more productive investment than procuring specialized digital tools, since the effect here derived primarily from pedagogical sequencing (contextual problem, collaborative investigation, gamified practice) rather than from the platform itself.

These conclusions should be interpreted in light of several limitations. First, the study was conducted in a single school with intact, non-randomly assigned classes; this quasi-experimental design limits internal validity relative to true experiments and constrains generalizability beyond similar settings. Second, the intervention spanned a short instructional period, and no delayed posttest was administered, so the durability of the observed motivational and cognitive gains beyond this short period remains untested. Third, the observer could not be blinded to group assignment, so some degree of observer-expectancy bias in fidelity and engagement ratings cannot be ruled out. Fourth, because CTL and gamification were implemented as a single combined treatment, their independent contributions to the observed effects cannot be disentangled within the present design.

Future research should therefore: (1) replicate this design across multiple schools and regencies with randomized assignment where feasible, to strengthen internal and external validity; (2) incorporate longitudinal follow-up assessments to test the durability of motivational and cognitive gains; (3) employ factorial designs that isolate the unique and combined contributions of CTL and gamification; (4) extend the intervention to other mathematical topics (e.g., linear equations, statistics) and grade levels to examine transferability; and (5) investigate the differential contribution of specific gamification features (leaderboards, badges, timers) to identify which design elements most benefit different learner profiles.

Acknowledgements
The authors express sincere gratitude to the principal, teachers, and students of SMP Negeri 2 Kaloran, Temanggung Regency, for their cooperation and participation in this study, and to Rohsepti Nurjannah, S.Pd. for serving as classroom observer. Special thanks are extended to the supervisors and faculty of the Master of Mathematics Education Program, School of Graduate Studies, Universitas Terbuka, for their invaluable guidance.
Declaration of Conflicting Interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.
Funding
This research was conducted as part of the first author's master's thesis research at the Master's Program in Mathematics Education, Universitas Terbuka, and received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Authors' contributions
N.N. (Nurmiati Nurmiati): conceptualization, fieldwork and data curation, formal analysis, investigation, writing – original draft; E.W. (Endang Wahyuningrum): methodology, validation, writing – review & editing, supervision; S.S. (Sudirman Sudirman): methodology, validation, writing – review & editing, project administration, supervision. All authors reviewed and approved the final manuscript.

Data availability
The data that support the findings of this study, including student test scripts, questionnaire responses, and classroom observation records, are not publicly available due to the personal nature of information shared by research participants, but are available from the corresponding author upon reasonable request.

Use of Artificial Intelligence

The authors declare that a generative artificial intelligence tool (Claude, Anthropic) was used to assist with proofreading and with compiling and formatting the bibliography. All research design, fieldwork, data analysis, interpretation, and conclusions are the sole intellectual responsibility of the authors, who reviewed, verified, and take full responsibility for the accuracy and integrity of the final manuscript.

6. References

Adeoye, B. F. (2024). Pedagogical models for contextual mathematics learning: A narrative synthesis. *International Journal of Educational Research Open*, *5*, 100279. https://doi.org/10.1016/j.ijedro.2023.100279

Akmal, M., & Pritchett, L. (2021). Learning equity requires more than equality: Learning goals and achievement gaps between the rich and the poor in five developing countries. *International Journal of Educational Development*, *82*, 102350. https://doi.org/10.1016/j.ijedudev.2021.102350

Almeqdad, Q. I., Alodat, A. M., Alquraan, M. F., Mohaidat, M. A., & Al-Makhzoomy, A. K. (2023). The effectiveness of universal design for learning: A systematic review of the literature and meta-analysis. *Cogent Education*, *10*(1), 2218191. https://doi.org/10.1080/2331186X.2023.2218191

Bai, S., Hew, K. F., & Huang, B. (2020). Does gamification improve student learning outcome? Evidence from a meta-analysis and synthesis of qualitative data in educational contexts. *Educational Research Review*, *30*, 100322. https://doi.org/10.1016/j.edurev.2020.100322

Bandura, A. (1997). Self-efficacy: The exercise of control. W. H. Freeman.

Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. In M. A. Gernsbacher et al. (Eds.), Psychology and the real world (pp. 56–64). Worth Publishers.

Boström, E., & Palm, T. (2023). The effect of a formative assessment practice on student achievement in mathematics. *Frontiers in Education*, *8*, 1101192. https://doi.org/10.3389/feduc.2023.1101192

Businskas, A. M. (2008). Conversations about connections: How secondary mathematics teachers conceptualize and contend with mathematical connections. https://summit.sfu.ca/_flysystem/fedora/2025-07/etd4202a.pdf

Camacho-Sánchez, R., Serna Bardavío, J., Rillo-Albert, A., & Lavega-Burgués, P. (2023). Enhancing motivation and academic performance through gamified digital game-based learning methodology using the ARCS model. *Interactive Learning Environments*, *32*(10), 6868–6885. https://doi.org/10.1080/10494820.2023.2294762

Cheung, A. C. K., & Slavin, R. E. (2013). The effectiveness of educational technology applications for enhancing mathematics achievement in K-12 classrooms: A meta-analysis. *Educational Research Review*, *9*, 88–113. https://doi.org/10.1016/j.edurev.2013.01.001

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Lawrence Erlbaum Associates.

Creswell, J. W. (2012). Educational research: Planning, conducting, and evaluating quantitative and qualitative research (4th ed.). Pearson.

Deci, E. L., & Ryan, R. M. (1985). Intrinsic motivation and self-determination in human behavior. Plenum Press.

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness: Defining gamification. In Proceedings of the 15th International Academic MindTrek Conference (pp. 9–15). ACM. https://doi.org/10.1145/2181037.2181040

Dolores-Flores, C., Rivera-López, M. I., & García-García, J. (2019). Exploring mathematical connections of pre-university students through tasks involving rates of change. *International Journal of Mathematical Education in Science and Technology, 50*(3), 369–389. https://doi.org/10.1080/0020739X.2018.1507050

Freudenthal, H. (1991). Revisiting mathematics education: China lectures. Kluwer Academic.

García-García, J., & Dolores-Flores, C. (2021). Exploring pre-university students' mathematical connections when solving Calculus application problems. *International Journal of Mathematical Education in Science and Technology, 52*(6), 912–936. https://doi.org/10.1080/0020739X.2020.1729429

Gust, S., Hanushek, E. A., & Woessmann, L. (2024). Global universal basic skills: Current deficits and implications for world development. *Journal of Development Economics, 166*, 103205. https://doi.org/10.1016/j.jdeveco.2023.103205

Gyenyn, M., Adu-Obeng, B., & Arthur, Y. D. (2026). Guided teaching, technology integration, and learner engagement as predictors of mathematical problem-solving ability: A structural equation modeling study. *International Journal of Didactic Mathematics in Distance Education, 3*(1), 119–136. Retrieved from https://jurnal.ut.ac.id/index.php/ijdmde/article/view/13962

Hatisaru, V. (2024). Mathematical connections – a growing construct. *International Journal of Mathematical Education in Science and Technology, 55*(3), 585–589. https://doi.org/10.1080/0020739X.2023.2283346

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112. https://doi.org/10.3102/003465430298487

Herset, M., El Ghami, M., & Bjerke, A. H. (2023). The effect of level-marked mathematics tasks on students' self-efficacy: An experimental study. *Frontiers in Psychology, 14*, 1116386. https://doi.org/10.3389/fpsyg.2023.1116386

Hofer, S. I., & Reinhold, F. (2025). Scaffolding of learning activities: Aptitude-treatment-interaction effects in math? *Learning and Instruction, 99*, 102177. https://doi.org/10.1016/j.learninstruc.2025.102177

Howard, J. L., Bureau, J., Guay, F., Chong, J. X. Y., & Ryan, R. M. (2021). Student motivation and associated outcomes: A meta-analysis from self-determination theory. *Perspectives on Psychological Science, 16*(6), 1300–1323. https://doi.org/10.1177/1745691620966789

Huang, R., Ritzhaupt, A. D., Sommer, M., Zhu, J., Stephen, A., Valle, N., Hampton, J., & Li, J. (2020). The impact of gamification in educational settings on student learning outcomes: A meta-analysis. *Educational Technology Research and Development, 68*(4), 1875–1901. https://doi.org/10.1007/s11423-020-09807-z

Johnson, E. B. (2002). Contextual teaching and learning: What it is and why it's here to stay. Corwin Press.

Kang, H., & Cogan, L. (2022). The differential role of socioeconomic status in the relationship between curriculum-based mathematics and mathematics literacy: The link between TIMSS and PISA. *International Journal of Science and Mathematics Education, 20*(1), 133–148. https://doi.org/10.1007/s10763-020-10133-2

Kaya, O. S., & Ercag, E. (2023). The impact of applying challenge-based gamification program on students' learning outcomes: Academic achievement, motivation and flow. *Education and Information Technologies, 28*, 10053–10078. https://doi.org/10.1007/s10639-023-11585-z

Keller, J. M. (2010). Motivational design for learning and performance: The ARCS model approach. Springer.

Kenedi, A. K., Helsa, Y., Ariani, Y., Zainil, M., & Hendri, S. (2019). Mathematical connection of elementary school students to solve mathematical problems. *Journal on Mathematics Education*, *10*(1), 69–80. https://doi.org/10.22342/jme.10.1.5416.69-80

Khairunnisak, C., Johar, R., Maulina, S., Zubainur, C. M., & Maidiyah, E. (2024). Teachers' understanding of realistic mathematics education through a blended professional development workshop on designing learning trajectory. *International Journal of Mathematical Education in Science and Technology*, *55*(4), 805–828. https://doi.org/10.1080/0020739X.2022.2038800

Leroy, N., & Bressoux, P. (2016). Does amotivation matter more than motivation in predicting mathematics learning gains? A longitudinal study of sixth-grade students in France. *Contemporary Educational Psychology*, *44–45*, 41–53. https://doi.org/10.1016/j.cedpsych.2015.11.003

Li, C., Liang, L., Fryer, L. K., & Shum, A. (2024). The use of leaderboards in education: A systematic review of empirical evidence in higher education. Journal of Computer Assisted Learning, 40(6), 3406–3442. https://doi.org/10.1111/jcal.13077

Li, L., Hew, K. F., & Du, J. (2024). Gamification enhances student intrinsic motivation, perceptions of autonomy and relatedness, but minimal impact on competency: A meta-analysis and systematic review. *Educational Technology Research and Development*, *72*(2), 765–796. https://doi.org/10.1007/s11423-023-10337-7

Li, M., Ma, S., & Shi, Y. (2023). Examining the effectiveness of gamification as a tool promoting teaching and learning in educational settings: A meta-analysis. *Frontiers in psychology*, *14*, 1253549. https://doi.org/10.3389/fpsyg.2023.1253549

Li, Q., & Ma, X. (2010). A meta-analysis of the effects of computer technology on school students' mathematics learning. *Educational Psychology Review*, *22*(3), 215–243. https://doi.org/10.1007/s10648-010-9125-8

Li, Y., Jiang, C., Chen, Z., et al. (2023). Peer tutoring models in collaborative learning of mathematical problem solving and their effect on group achievement. *Education and Information Technologies*, *28*, 6595–6618. https://doi.org/10.1007/s10639-022-11429-2

Liu, X., Yang-Hansen, K., Valcke, M., & De Neve, J. (2024). A decade of PISA: Student perceived instructional quality and mathematics achievement across European countries. *ZDM – Mathematics Education*, *56*(5), 859–891. https://doi.org/10.1007/s11858-024-01630-7

Luarn, P., Chen, C.-C., & Chiu, Y.-P. (2023). Enhancing intrinsic learning motivation through gamification. *International Journal of Information and Learning Technology*, *40*(5), 413–424. https://doi.org/10.1108/IJILT-07-2022-0145

Lyle, K. B., Bego, C. R., Ralston, P. A. S., & Immekus, J. C. (2022). Spaced retrieval practice imposes desirable difficulty in calculus learning. *Educational Psychology Review*, *34*(3), 1799–1812. https://doi.org/10.1007/s10648-022-09677-2

Mayangsari, C., Qohar, A., & Faridha, N. (2025). Enhancing mathematical disposition and learning outcomes through Team Games Tournament: A two-cycle action research on probability instruction in indonesian secondary education. *Polyhedron International Journal in Mathematics Education*, *3*(2), 110–126. https://doi.org/10.59965/pijme.v3i2.205

Muzaki, A., Juliangkary, E., & Zaitun. (2023). Pengaruh pembelajaran kontekstual terhadap kemampuan koneksi matematis siswa [The effect of contextual learning on students' mathematical connection ability]. *Mandalika Mathematics and Education Journal*, *5*(1), 60–70. https://doi.org/10.29303/jm.v5i1.4048

National Council of Teachers of Mathematics (NCTM). (2000). Principles and standards for school mathematics. NCTM.

Ningsih, R. (2021). Model CTL dengan bahan ajar gamifikasi terhadap kemampuan pemahaman konsep matematis [The effect of the contextual teaching and learning model with gamified teaching materials on students' mathematical conceptual understanding]. *Jurnal Cendekia: Jurnal Pendidikan Matematika*, *5*(2), 1283–1294. https://doi.org/10.31004/cendek ia.v5i2.724

OECD. (2023). PISA 2022 results (Volume I): The state of learning and equity in education. OECD Publishing. https://doi.org/10.1787/53f23881-en

Ortiz-Rojas, M., Chiluiza, K., Valcke, M., & Bolaños-Mendoza, C. (2025). How gamification boosts learning in STEM higher education: A mixed methods study. *International Journal of STEM Education*, *12*(1), 1–14. https://doi.org/10.1186/s40594-024-00521-3

Pellegrini, M., Lake, C., Neitzel, A., & Slavin, R. E. (2021). Effective programs in elementary mathematics: A meta-analysis. *AERA Open*, *7*. https://doi.org/10.1177/2332858420986211

Pintrich, P. R., & Schunk, D. H. (2002). Motivation in education: Theory, research, and applications (2nd ed.). Merrill Prentice Hall.

Putri, L. I., Retnawati, H., & Jaedun, A. (2025). Enhancing mathematical skills through multicontextual approaches: A meta-analysis of realistic mathematics, ethnomathematics, and technology integration. *Cogent Education*, *12*(1), 2548648. https://doi.org/10.1080/2331186X.2025.2548648

Rodríguez-Nieto, C. A., Cabrales, H. A., Arenas-Peñaloza, J., Schnorr, V., & Font, V. (2024). Onto-semiotic analysis of Colombian engineering students' mathematical connections to problem-solving on vectors: A contribution to the natural and exact sciences. *EURASIA Journal of Mathematics, Science and Technology Education*, *20*(5), 2–24. https://doi.org/10.29333/ejmste/14450

Rodríguez-Nieto, C. A., Rodríguez-Vásquez, F. M., & Font, V. (2022). A new view about connections: The mathematical connections established by a teacher when teaching the derivative. *International Journal of Mathematical Education in Science and Technology*, *53*(6), 1231–1256. https://doi.org/10.1080/0020739X.2020.1799254

Rodríguez-Nieto, C. A., Rodríguez-Vásquez, F. M., & Font, V. (2023). Combined use of the extended theory of connections and the onto-semiotic approach to analyze mathematical connections by relating the graphs of f and f'. *Educational Studies in Mathematics*, *114*(1), 63–88. https://doi.org/10.1007/s10649-023-10246-9

Rose, D. H., & Meyer, A. (2002). Teaching every student in the digital age: Universal Design for Learning. ASCD.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation. *American Psychologist*, *55*(1), 68–78. https://doi.org/10.1037/0003-066X.55.1.68

Sailer, M., & Homner, L. (2020). The gamification of learning: A meta-analysis. *Educational Psychology Review*, *32*(1), 77–112. https://doi.org/10.1007/s10648-019-09498-w

Sardiman, A. M. (2014). Interaksi dan motivasi belajar mengajar [Interaction and Motivation in Teaching and Learning]. Rajawali Press.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). Experimental and quasi-experimental designs for generalized causal inference. Houghton Mifflin.

Söderström, S., & Palm, T. (2024). Feedback in mathematics education research: A systematic literature review. *Research in Mathematics Education*, *28*(1), 163–184. https://doi.org/10.1080/14794802.2024.2401488

Tomlinson, C. A. (2001). How to differentiate instruction in mixed-ability classrooms (2nd ed.). ASCD.

van Balen, J., Gosen, M., de Vries, S., & Koole, T. (2024). Peer-to-peer-talk in whole-classroom discussions. *International Journal of Educational Research*, *125*, 102354. https://doi.org/10.1016/j.ijer.2024.102354

Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes. Harvard University Press.

Wang, X. F., Perry, L. B., Malpique, A., & Ide, T. (2023). Factors predicting mathematics achievement in PISA: A systematic review. *Large-Scale Assessments in Education*, *11*, 24. https://doi.org/10.1186/s40536-023-00174-8

Wardani, E. K., Retnasari, L. ., & Ghiffari, M. A. N. (2025). AR Snakes and ladders game innovation to romote learning to live together in elementary schools. *International Journal of Mathematics and Sciences Education*, *3*(2), 57–70. https://doi.org/10.59965/ijmsed.v3i2.264

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